Protein - Protein Interaction Studies by Chemical Cross-Linking and Mass Spectrometry

D i s s e r t a t i o n
zur Erlangung des akademischen Grades
Dr. rer. nat.

vorgelegt der
Naturwissenschaftlichen Fakultät I
Biowissenschaften
der Martin-Luther-Universität Halle-Wittenberg

von
Daniela M. Schulz
geboren am 28. Juli 1978 in Nürnberg

Gutachter

1. Prof. Dr. Andrea Sinz

2. Prof. Dr. Milton Stubbs

3. Prof. Dr. Michael Glocker

Eingereicht am: 14.02.2007
Tag der öffentlichen Verteidigung: 24.07.2007

Für Oma und Opa

Content

Bibliographische Beschreibung

Daniela M. Schulz

Protein - Protein Interaction Studies by Chemical Cross-Linking and Mass Spectrometry

A critical step towards gaining deeper insights into protein / protein interactions is the elucidation of the spatial organization of protein complexes. Chemical cross-linking - the process of covalent attachment of near-neighbor amino acid functional groups - in combination with high-resolution mass spectrometry constitutes a powerful analytical approach for the structural characterization of protein complexes. The interacting regions within the 20 kDa-complex between calmodulin (CaM) and melittin, and thus the orientation of melittin within the complex were determined. Covalent conjugation of the interacting proteins was achieved by applying different cross-linking reagents, and both MALDI-TOFMS and SDS-PAGE of the reaction mixtures evidenced the formation of a cross-linked complex between CaM and melittin. Screening of the mass spectra obtained from nano-HPLC/nano-ESI-FTICRMS of enzymatically digested cross-linked complexes led to the identification of 17 unique intermolecular cross-linked products. Evaluation of the data revealed not all of the obtained cross-links being consistent with a single orientation of melittin within the complex. The distance restraints derived from the chemical cross-linking data were used for generating low-resolution three-dimensional structure models for the CaM / melittin complex by computational methods. It was evidenced that CaM is able to recognize target peptides in two opposite orientations.

For structurally characterizing the heterotetrameric complex between annexin A2 (ANXA2) and S100A10 (p11) by chemical cross-linking and mass spectrometry, the complex was purified from mucosa of pig (*Sus scrofa*) small intestines. The ANXA2 / p11 complex (A2t) possesses a molecular weight of ~ 100 kDa and represents one of the largest protein complexes studied so far by chemical cross-linking and mass spectrometry. So far, no complete high-resolution annexin A2 / p11 structure is available. The spatial organization of the annexin A2 / p11 heterotetramer was scrutinized employing 1:1 mixtures of stable isotope-labeled, amine-reactive cross-linking reagents. This offers the decisive advantage of a facilitated identification of cross-linked products based on their characteristic isotope patterns. Furthermore, tandem mass spectrometry (MS/MS) proved invaluable for an exact localization of cross-linked amino acid residues. Based on the distance restraints obtained from the cross-linking data, structural models of the annexin A2 / p11 complex were derived by computational protein-protein docking methods, revealing an octameric conformation of the complex that differs from so far suggested models and provides new insight into annexin A2 / p11 interaction.

The identification of protein interaction partners allows to gain knowledge of networks in the cell. In this regard, chemical cross-linking represents a useful means through covalent attachment of interacting proteins, thus making even transient interactions observable. An approach for the identification of A2t interaction partners employing chemical cross-linking techniques was established, involving biotinylation of A2t, a two-step cross-linking procedure using a heterobifunctional amine- and photo-reactive cross-linker, and finally affinity purification on avidin beads of cross-linked A2t / interaction partner complexes via the biotin moiety and subsequent identification of interaction partners by mass spectrometric peptide mass fingerprint analysis.

Index of Tables and Figures

Tables

Figures

Abbreviations

Å	Ångstrom
aa	Amino acid
amu	Atomic mass unit
ANX	Annexin
APS	Ammoniumpersulfate
BS^{2}G	*Bis*(sulfosuccinimidyl) glutarate
BS^3	*Bis*(sulfosuccinimidyl) suberate
C20W	Synthetic peptide of the plasma membrane calcium pump
CA	C_α atom
CAD / CID	Collisionally-activated dissociation / collision-induced dissociation
CaM	Calmodulin
CaMK	CaM-dependent protein kinase
CaMKK	CaM-dependent kinase kinase
CO-1	Channel outlet 1
CO-2	Channel outlet 2
Da	Dalton
DE	Delayed extraction
DEAE	Diethylaminoethyl cellulose
DHB	2,5-Dihydroxy benzoic acid
DMSO	Dimethyl sulfoxide
DSA	Disuccinimidyl adipate
DTT	1,4-Dithiothreitol
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride
e.g.	for example
EGS	Ethylene glycol *bis*(succinimidly succinate)
ESI	Electrospray-ionization
FPLC	Fast protein liquid chromatography
FTICR	Fourier transform ion cyclotron resonance
HEPES	[4-(2-Hydroxyethyl)-piperazino]-ethane sulfonic acid
HPLC	High-performance liquid chromatography
i.e.	That is
kDa	Kilodalton
LC	Liquid chromatography
LHRH	Luteinizing hormone releasing hormone
MALDI	Matrix-assisted laser desorption/ionization
MARCKS	Myristoylated alanine-rich C kinase substrate
M13	Skeletal muscle myosin light chain kinase peptide
Mel	Melittin
MES	2-[*N*-Morpholino]ethane sulfonic acid
MLCK	Myosin light chain kinase
MS	Mass spectrometer, Mass spectrometry
MS/MS	Tandem mass spectrometry

m/z	Mass to charge ratio	
n.a.	not assigned	
NHS	*N*-Hydroxysuccinimide	
NZ	Nitrogen atom of the lysine sidechain	
OG	Oxygen atom of the serine sidechain	
p11	S100A10, 11kDa protein of the S100 protein family	
PAGE	Polyacrylamide gel electrophoresis	
PDB	Protein Data Bank of the Research Collaboratory for Structural Bioinformatics	
pI	Isoelectric point	
PMF	Peptide mass fingerprint	
ppm	Parts per million	
R	Resolution	
RP	Reversed-phase	
S100A10	*see* p11	
SANPAH	*N*-Succinimidyl-6-[4'-azido-2'-nitrophenylamino] hexanoate	
SDS	Sodium dodecylsulfate	
sDST	Disulfosuccinimidyl tartrate	
sNHS	*N*-Hydroxysulfosuccinimide	
skMLCK	Skeletal muscle myosin light chain kinase	
smMLCK	Smooth muscle myosin light chain kinase	
TEMED	*N,N,N´,N´*-Tetramethylethylenediamine	
TFA	Trifluoro acetic acid	
TIC	Total ion current	
TOF	Time-of-flight	
TRIS	Tris-(hydroxymethyl)-aminomethane	
u	*see* amu	
v/v	Volume / volume	
w/v	Weight / volume	

Proteinogenic amino acids

Alanine	Ala	**A**	Leucine	Leu	**L**
Arginine	Arg	**R**	Lysine	Lys	**K**
Asparagine	Asn	**N**	Methionine	Met	**M**
Aspartic acid	Asp	**D**	Phenylalanine	Phe	**F**
Cysteine	Cys	**C**	Proline	Pro	**P**
Glutamine	Gln	**Q**	Serine	Ser	**S**
Glutamic acid	Glu	**E**	Threonine	Thr	**T**
Glycine	Gly	**G**	Tryptophan	Trp	**W**
Histidine	His	**H**	Tyrosine	Tyr	**Y**
Isoleucine	Ile	**I**	Valine	Val	**V**

Zusammenfassung

Gegenwärtig stellt die Aufklärung von Protein-Protein Interaktionen eine der wesentlichen Herausforderungen auf dem Gebiet der Proteinanalytik dar. Damit eng verbunden ist die Aufklärung der Funktion von Proteinen und deren Zusammenspiel in der Zelle. Um die Funktionsweisen von Proteinen verstehen zu können, ist es wichtig, die zugrunde liegenden strukturellen Zusammenhänge zu kennen. Bisher war die Aufklärung der räumlichen Struktur von Proteinen und Proteinkomplexen den etablierten, hochauflösenden Methoden, wie der Röntgenstrukturanalyse und der NMR (Kernmagnetresonanz)-Spektroskopie vorbehalten. Jedoch ermöglichten die in den letzten Jahren erzielten Fortschritte auf dem Gebiet der Massenspektrometrie die Entwicklung alternativer experimenteller Ansätze für die Strukturaufklärung von Proteinen und Proteinkomplexen.

Chemisches Cross-Linking[1] - eine Methode zur kovalenten Verknüpfung räumlich benachbarter funktioneller Gruppen von Aminosäuren mittels eines chemischen Reagenzes - stellt in Verbindung mit massenspektrometrischen Methoden einen analytischen Gegenpart zur Röntgenstrukturanalyse und zur NMR-Spektroskopie dar, der die Bestimmung niederaufgelöster Proteinstrukturen ermöglicht. Der auf chemischem Cross-Linking und Massenspektrometrie basierende Ansatz erreicht zwar nicht die Hochauflösung, die mit Röntgenstrukturanalyse und NMR-Spektroskopie erhalten wird, jedoch weist er einige entscheidende Vorteile auf: Die Experimente sind vergleichsweise schnell durchzuführen, es werden nur sehr geringe (im subfemtomolaren Bereich liegende) Analytmengen benötigt und die Größe der zu untersuchenden Proteine ist theoretisch unbegrenzt, da die enzymatischen Spaltprodukte analysiert werden. Des Weiteren eröffnen die große Vielfalt an Cross-Linking-Reagenzien mit unterschiedlichen Spezifitäten für bestimmte funktionelle Gruppen, wie beispielsweise für primäre Amin- oder Carboxylgruppen, und die unterschiedlichen Längen von Cross-Linking-Reagenzien zahlreiche Möglichkeiten des experimentellen Designs.

In dieser Arbeit wurden chemisches Cross-Linking und Massenspektrometrie für die Strukturaufklärung des 20 kDa-Calmodulin / Melittin-Komplexes eingesetzt und als Weiterentwicklung auf den 100 kDa-Heterotetramer-Komplex zwischen Annexin A2 (ANXA2) und S100A10 (p11) angewendet. Der ANXA2 / p11-Komplex (A2t) wurde aus der Mucosa des Schweinedünndarms (*Sus scrofa*) isoliert und gereinigt. Darüber hinaus wurde ein Protokoll auf

---

[1] Es wird der englische Ausdruck „Cross-Linking", anstatt „Quervernetzung" verwendet, da sich dieser in der Literatur durchgesetzt hat.

Grundlage von Cross-Linking und Massenspektrometrie zur Identifizierung spezifischer A2t Bindungspartner entwickelt.

Die auf der Kombination von chemischem Cross-Linking und Massenspektrometrie basierende analytische Strategie zur Untersuchung der Struktur von Proteinkomplexen mittels intermolekularen Cross-Linkings beginnt mit der kovalenten Verknüpfung funktioneller Gruppen von Aminosäureresten, die sich enger räumlicher Nachbarschaft zueinander befinden. Hierfür steht eine Vielzahl an Reagenzien zur Verfügung. Die entstandenen Reaktionsgemische der intakten vernetzten Proteinkomplexe werden mit Hilfe der MALDI-TOFMS im Hinblick auf das Ausmaß der Cross-Linking-Reaktion untersucht. Die Reaktionsgemische werden mit eindimensionaler Gelelektrophorese aufgetrennt, wobei nicht verknüpfte Untereinheiten des Komplexes dissoziieren und so von den höhermolekularen, verknüpften Proteinkomplexen getrennt werden. Die Gelbanden, die für die nachfolgende Analyse von Interesse sind, werden aus dem Gel ausgeschnitten und enzymatisch gespalten. Dies führt zur Bildung einer hochkomplexen Mischung bestehend aus unmodifizierten Peptiden und intramolekular (intrapeptidal oder interpeptidal) verknüpften Peptiden, sowie, im günstigsten Fall, aus intermolekularen Cross-Linking Produkten. Weiterhin besteht die Möglichkeit der Modifikation von Peptiden mit einem partiell hydrolysierten Cross-Linker, bei Verwendung von Cross-Linkern, die eine *N*-Hydroxysuccinimidgruppe enthalten. Die erhaltenen Peptidmischungen werden massenspektrometrisch analysiert und die Zuordnung der Cross-Linking Produkte erfolgt anhand des Vergleichs der erhaltenen *m/z*-Werte mit den theoretischen *m/z*-Werten der Peptide und Cross-Linking Produkte. Dadurch werden die Interaktionsflächen innerhalb des Komplexes bestimmt und die ungefähren Abstände der in die Interaktion einbezogenen Aminosäureseitenketten können so abgeschätzt werden.

Mit Hilfe der oben beschriebenen Strategie wurden die Kontaktregionen des Ca^{2+}-abhängigen Komplexes zwischen Calmodulin (CaM) und Melittin und damit gleichzeitig die Orientierung von Melittin innerhalb des Komplexes bestimmt. CaM und Melittin wurden vor der Untersuchung der räumlichen Struktur sehr genau bezüglich Aminosäuresequenz und posttranslationaler Modifizierungen charakterisiert. Zur kovalenten Verknüpfung der Proteine wurde EDC (1-Ethyl-3-(3-dimethylaminopropyl) carbodiimid) in Kombination mit sNHS (*N*-Hydroxysulfosuccinimid) eingesetzt, wodurch eine Amidbindung zwischen einer Amino- und einer Carboxylgruppe gebildet wird. Zusätzlich wurden die homobifunktionellen aminreaktiven Reagenzien sDST (Disulfosuccinimidyltartrat) und BS[3] (*Bis*(sulfosuccinimidyl)suberat) mit *Spacer*-Längen von 6.4 und 11.4 Ångström eingesetzt. Reagenzienmengen und Inkubationszeiten wurden optimiert um einerseits eine hohe Ausbeute des vernetzten Komplexes zu erhalten, gleichzeitig aber die Bildung hochmolekularer oligomerer Aggregate zu vermeiden. Sowohl MALDI-TOFMS und

2

SDS-PAGE der Reaktionsmischung bestätigten die Bildung des kovalent verknüpften CaM / Melittin Komplexes. Die enzymatischen Spaltprodukte der verknüpften Komplexe wurden mittels Nano-HPLC/Nano-ESI-FTICRMS analysiert. Die Auswertung der Massenspektren führte zur Identifizierung von 17 unterschiedlichen Cross-Linking Produkten, die jedoch nicht alle einer bestimmten Orientierung von Melittin innerhalb des Komplexes entsprachen, sondern vielmehr auf eine weitere Möglichkeit der Bindungsorientierung hindeuteten. Basierend auf den durch chemisches Cross-Linking erhaltenen Distanzbeschränkungen wurden mittels computergestützter Methoden des "*conjoined rigid body / torsion angle simulated annealing*" niederaufgelöste Strukturmodelle des CaM / Melittin Komplexes errechnet, für den bislang keine hochaufgelöste Struktur erhältlich ist. Es konnte gezeigt werden, dass CaM in der Lage ist, Zielpeptide in zwei entgegengesetzten Richtungen zu binden.

Für die Strukturaufklärung von A2t mittels chemischen Cross-Linkings und Massenspektrometrie, wurde der heterotetramere Komplex aus der Mucosa von Schweinedünndarm gereinigt und ANXA2 und p11 wurden bezüglich Molekulargewicht, Aminosäuresequenz und möglicher posttranslationaler Modifizierungen mittels Peptidmassenfingerprintanalyse untersucht. Die vollständige Aminosäuresequenz von ANXA2 (Schwein), für die in der Swiss-Prot-Datenbank nur das N-terminale Fragment bestehend aus Aminsäuren 1-91 erhältlich ist, konnte durch Vergleich mit den ANXA2 Einträgen anderer Säugetierspezies ermittelt werden. Die vollständige porcine ANXA2-Aminosäuresequenz wurde der Swiss-Prot-Datenbank inzwischen übermittelt. Für die erfolgreiche Durchführung von Strukturuntersuchungen mittels chemischen Cross-Linkings und Massenspektrometrie sind die Verfügbarkeit reinen Proteins sowie eine genaue Beschreibung der Zusammensetzung und Eigenschaften des Proteins unerlässlich.

Die vielfältigen Funktionen von A2t, das vor allem als Ca^{2+} / Phospholipid-bindendes Protein bekannt ist, und seine Bedeutung in einer ganzen Reihe von Krankheiten, macht A2t zu einem interessanten Studienobjekt. Außerdem ist A2t mit einem Molekulargewicht von ~ 100 kDa einer der größten Komplexe, die jemals mit chemischem Cross-Linking und Massenspektrometrie untersucht wurden. Wenngleich hochaufgelöste Strukturen für die beiden interagierenden Proteine und für den Komplex zwischen dem p11-Dimer und zwei N-terminalen ANXA2-Peptiden vorliegen, so wurde die vollständige Struktur des Heterotetramers bislang noch nicht gelöst. Als biologisch relevante Komplexstrukturen des A2t werden sowohl die heterotetramere als auch die oktamere Form diskutiert.

In dieser Arbeit wurde die dreidimensionale Struktur des Annexin A2 / p11-Tetramers (A2t) untersucht. Hierfür wurden die homobifunktionellen Cross-Linking-Reagenzien sDST (Disulfosuccinimidyltartrat), BS^3-d_0/d_4 (*Bis*(sulfosuccinimidyl)suberat), BS^2G-d_0/d_4 (*Bis*(sulfosuccinimidyl)glutarat), and DSA-d_0/d_8 (Disuccinimidyladipat) mit *Spacer*-Längen

zwischen 6.4 und 11.4 Ångström eingesetzt. Die drei letztgenannten Reagenzien wurden als 1:1 Mischungen ihrer nicht-deuterierten (d_0) and deuterierten (d_4 oder d_8) Derivate eingesetzt, was die Identifizierung von Cross-Linking-Produkten aufgrund der 4 (d_0/d_4) oder 8 (d_0/d_8) amu-Abstände ihrer jeweiligen Isotopenmuster der einfach geladenen Signale in MALDI-TOF-Massenspektren und dekonvolutierten ESI-FTICR-Massenspektren entscheidend vereinfacht. Des Weiteren erwies sich die Tandem-Massenspektrometrie (MS/MS) als außerordentlich wichtig für die präzise Lokalisierung der verknüpften Aminosäuren sowie für die eindeutige Zuordnung von Cross-Linking-Produkten. Die erhaltenen Distanzbeschränkungen wurden für die Berechnung von Modellen des ANXA2 / p11-Komplexes mit computergestützten Protein-Protein-Docking Programmen verwendet. Die Strukturmodelle des ANXA2 / p11-Komplexes offenbarten eine oktamere Konformation. Die errechnete Oktamerstruktur unterscheidet sich von bislang vorgeschlagenen Modellen und eröffnet neue Einblicke die Interaktion zwischen Annexin A2 und p11.

Die Bestimmung von Proteininteraktionspartnern erfreut sich wachsenden Interesses mit dem Ziel, wichtige Einblicke in Proteinnetzwerke zu gewinnen und Signaltransduktionswege aufzuklären. Für das Erreichen dieses Ziels stellt chemisches Cross-Linking ein wertvolles Mittel dar, da durch die kovalente Verknüpfung von Interaktionspartnern auch transiente Interaktionen sichtbar gemacht werden können.

Deshalb war ein weiteres Ziel dieser Arbeit, einen auf chemischem Cross-Linking basierten Ansatz für die Identifizierung von A2t-Interaktionspartnern zu entwickeln. Das entwickelte Protokoll umfasst die Biotinylierung von A2t, eine Zweistufenstrategie für das Cross-Linking mit einem heterobifunktionellen, amin- und photoreaktiven Reagenz und schließlich die spezifische Anreicherung der verknüpften A2t / Interaktionspartner-Komplexe durch Affinitätsreinigung der biotinylierten Komplexe an immobilisiertem Avidin. Die vermeintlichen Interaktionspartner wurden per MALDI-TOFMS-Peptidmassenfingerprintanalyse identifiziert.

In der vorliegenden Arbeit konnte gezeigt werden, dass die Kombination aus chemischem Cross-Linking, Massenspektrometrie und computergestützten Methoden einen leistungsstarken Ansatz für die strukturelle Charakterisierung darstellt. Der analytische Ansatz wurde erfolgreich am Beispiel eines Protein-Peptid- (CaM / Melittin) und eines Protein-Protein (ANXA2 / p11)-Komplexes eingesetzt.

Die Zuordnung der Cross-Linking-Produkte ist für relativ kleine Proteinkomplexe, wie dem 20 kDa-Komplex zwischen CaM und Melittin, allein anhand der hochgenauen Massen möglich. Es konnte hingegen gezeigt werden, dass für die korrekte Zuordnung der Cross-Linking-Produkte größerer Proteinkomplexe sowohl der Einsatz isotopenmarkierter Reagenzien, als auch der Tandem-Massenspektrometrie von außerordentlicher Wichtigkeit ist.

Summary

Currently, one major issue in protein science is deciphering protein / protein interactions, thus developing an understanding of protein function and of the protein interaction networks in the cell. A critical step towards gaining deeper insights into protein / protein interactions is the elucidation of the spatial organization of protein complexes. To date, X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy are unrivaled in obtaining high-resolution protein structures. However, owing to recent advances in mass spectrometric instrumentation, mass spectrometry-based approaches for the structural characterization of proteins and protein complexes have evolved as promising alternative.

Chemical cross-linking - the process of covalent attachment of near-neighbor amino acid functional groups - in combination with high-resolution mass spectrometry provides an analytical approach that can serve as a low-resolution counterpart to X-ray crystallography and NMR spectroscopy in deriving protein structures. While the combined chemical cross-linking / mass spectrometry approach is certainly not able to compete with the high-resolution structural data provided by both X-ray crystallography and NMR spectroscopy, it has some decisive advantages. The mass range of the protein complex under scrutiny is theoretically unlimited, since it is the proteolytic peptides that are analyzed. Analysis is generally fast and requires only minute - down to subfemtomole - amounts of analyte. Further, the broad range of specificities available for cross-linking reagents towards certain functional groups, such as primary amines or carboxylic acids, and the wide range of distances different cross-linking reagents can bridge, offer various options for experimental design.

In this work, chemical cross-linking and mass spectrometry were employed for structural characterization of the 20 kDa calmodulin / melittin complex and the established protocol was advanced to elucidating the quaternary structure of the 100 kDa heterotetramer complex between annexin A2 (ANXA2) and p11 (S100A10). The ANXA2 / p11 complex (A2t) was purified from small intestinal mucosa of the pig (*Sus scrofa*). Moreover, a protocol for the identification of A2t interaction partners based on cross-linking methods and mass spectrometric analysis was developed.

The analytical strategy of combined chemical cross-linking and mass spectrometric analysis for exploring protein complex structures by intermolecular cross-linking starts with the covalent attachment of functional groups of amino acid residues that are in close spatial proximity. For this purpose a wide variety of cross-linking reagents is available. The cross-linking reaction mixtures of the intact protein complex are analyzed by MALDI-TOFMS for monitoring the extent of cross-linking. The reacted protein complex is then separated by one-dimensional gel electrophoresis (SDS-PAGE). At this stage, the protein complex subunits that are not cross-

linked diverge and are thus separated from cross-linked species migrating at higher molecular weight. The gel bands of interest are excised and subjected to *in-gel* enzymatic proteolysis, which yields highly intricate mixtures of unmodified peptides, intramolecular intra- and interpeptide cross-linked products, and, most favorably, intermolecular cross-linked products. As cross-linkers containing an *N*-hydroxysuccinimide group are susceptible to hydrolysis, peptides modified with partially hydrolyzed cross-linkers are also observed. From the mass spectrometric analysis of the proteolytic peptide mixtures peaklist are generated and comparison of the obtained *m/z* values to the theoretical *m/z* values of peptides and cross-linked products allows assignment of cross-linked products. Thus, the interaction sites within a protein complex are revealed and one can estimate the approximate distance of these amino acid residues involved in the interaction.

Pursuing the strategy described above, the interacting regions within the calcium-dependent complex between calmodulin (CaM) and melittin, and thus the orientation of melittin within the complex were determined. Prior to topological investigation, CaM and melittin were characterized with respect to amino acid sequence and post-translational modifications. For covalent conjugation of the interacting proteins the zero-length cross-linkers EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride) and sNHS (*N*-hydroxysulfosuccinimide) that form an amide bond between an amine and a carboxylic acid was utilized. Additionally, the homobifunctional, amine-reactive cross-linking reagents sDST (disulfosuccinimidyl tartrate) and BS^3 (*bis*(sulfosuccinimidyl)suberate) bridging distances of 6.4 and 11.4 Angstrom, respectively, were employed. Amounts of cross-linking reagents and incubation times were optimized with respect to high cross-linking yield, but avoiding the formation of high-molecular weight oligomeric aggregates. Both MALDI-TOFMS and SDS-PAGE of the reaction mixtures evidenced the formation of a cross-linked complex between CaM and melittin. Enzymatically digested cross-linked complexes were analyzed by nano-HPLC/nano-ESI-FTICRMS. Screening of the obtained mass spectra led to the identification of 17 unique intermolecular cross-linked products. Evaluation of the data revealed not all of the obtained cross-links being consistent with a single orientation of melittin within the complex. Using the distance restraints derived from the chemical cross-linking data in combination with computational methods of conjoined rigid body / torsion angle simulated annealing, low-resolution three-dimensional structure models for the CaM / melittin complex were generated, for which no high-resolution structure exists to date. It was evidenced that CaM is able to recognize target peptides in two opposite orientations.

For structurally characterizing A2t by chemical cross-linking and mass spectrometry, the complex was purified from mucosa of pig (*Sus scrofa*) small intestines and the constituent

proteins were thoroughly scrutinized with respect to molecular weight, amino acid sequence and possibly present post-translational modifications by peptide mass fingerprint analysis. The complete amino acid sequence of porcine ANXA2, for which only the N-terminal fragment comprising amino acids 1-91 is deposited in the Swiss-Prot Database, was assessed by comparison with ANXA2 entries of other vertebrate species. The complete amino acid sequence of ANXA2 (pig) has in the meantime been submitted to the Swiss-Prot Database. The availability of pure protein and an accurate and detailed description of a protein's composition and properties are indispensable for a successful outcome of structural studies by chemical cross-linking and mass spectrometry.

The diverse biological functions of A2t, renowned as Ca^{2+} / phospholipid binding protein, and its implication in a number of diseases makes A2t an interesting protein complex to be studied. Furthermore, A2t possesses a molecular weight of ~ 100 kDa and thus constitutes one of the largest protein complexes studied so far by chemical cross-linking and mass spectrometry. Although high-resolution data are available for both interaction partners as well as for the complex between the p11 dimer and two annexin A2 N-terminal peptides, the structure of the complete annexin A2 / p11 heterotetramer has not yet been solved at high resolution. Thus, the quaternary structure of the biologically relevant, membrane-bound annexin A2 / p11 complex is still under discussion, while the existence of a heterotetramer or a heterooctamer is the prevailing opinion.

In this work, the spatial organization of the annexin A2 / p11 heterotetramer was scrutinized employing the homobifunctional cross-linkers sDST (disulfosuccinimidyl tartrate), BS^3-d_0/d_4 (bis(sulfosuccinimidyl) suberate), BS^2G-d_0/d_4 (bis(sulfosuccinimidyl) glutarate), and DSA-d_0/d_8 (disuccinimidyl adipate) bridging distances between 6.4 and 11.4 Angstrom. The latter three reagents were employed as 1:1 mixtures of their non-deuterated (d_0) and deuterated (d_4 or d_8) derivatives, which offers the decisive advantage of facilitated identification of cross-linked products from the distinct 4 (d_0/d_4) or 8 (d_0/d_8) amu spacing of the respective isotopic patterns of singly charged signals ($[M+H]^+$) obtained in MALDI-TOF mass spectra and deconvoluted ESI-FTICR mass spectra. Furthermore, tandem mass spectrometry (MS/MS) proved invaluable for an exact localization of cross-linked amino acid residues and for a confirmation of correct cross-linked product assignment. Based on the distance restraints obtained from the cross-linking data, structural models of the annexin A2 / p11 complex were derived by computational protein-protein docking methods, revealing an octameric conformation of the complex. The proposed structure of the annexin A2 / p11 octamer differs from so far suggested models and provides new insight into annexin A2 / p11 interaction.

The identification of protein interaction partners has attended increasing interest, providing important insights into protein networking and for elucidating signal transduction pathways. For

this very purpose, chemical cross-linking represents a useful means through covalent attachment of interacting proteins, thus making transient interactions observable.

Thus, another goal of this work was to establish an approach for the identification of A2t interaction partners employing chemical cross-linking techniques. A protocol was developed that involves biotinylation of A2t, a two-step cross-linking procedure using a heterobifunctional amine- and photo-reactive cross-linker, and finally affinity purification on avidin beads of cross-linked A2t / interaction partner complexes via the biotin moiety. The putative interactions partners were identified by MALDI-TOF mass spectrometric peptide mass fingerprint analysis.

In the present work, it was demonstrated that chemical cross-linking in conjunction with mass spectrometry and computational methods represents a powerful approach for the structural characterization of protein complexes. The strategy was successfully exemplified for a protein /peptide and a protein / protein complex, namely the calmodulin / melittin and the annexin A2 / p11 complex.

Cross-linked product assignment merely based on exact mass measurements proved feasible for relatively small protein complexes, such as the 20 kDa-complex between CaM and melittin, whereas for larger protein assemblies both the application of isotope-labeled cross-linking reagents and tandem mass spectrometry were indispensable for an unambiguous assignment of cross-linked products.

1 Theoretical Background

1.1 Structural Characterization of Proteins

With the human genome sequenced [International Human Genome Sequencing Consortium, 2001, Venter *et al.*, 2001], increasing attention is directed towards analyzing proteins. Knowledge of a protein's three-dimensional structure is critical for understanding its function. Numerous proteins exert their biological function through interactions with other proteins. Furthermore, proteins are the principal pharmaceutical targets. These facts explain why great efforts are being made for solving tertiary and quaternary protein structures and point out the importance of exploring the protein-protein interaction networks of the cell.

Nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography provide high-resolution protein structural data. By the end of 2006, 37,481 protein structures solved by NMR spectroscopy and X-ray crystallography were deposited in the Protein Data Bank (PDB) [www.rcsb.org/pdb], and every year an increasing number of several thousands of new structures are submitted. In the past ten years almost 100 protein structures solved by electron microscopy were added to the PDB, demonstrating the evolving relevance of this technique. Despite the impressive number of high-resolution structures available, the number of solved protein complex structures is comparably low (561, as of November 2006) [PROTCOM database: www.ces.clemson.edu/compbio/protcom/, Kundrotas & Alexov 2007], as is the number of membrane proteins (~ 100, [von Heijne, 2006]). The reason is closely connected to the inherent challenges of NMR spectroscopy and X-ray crystallography: Sufficient amounts of protein for analysis often are difficult to be obtained, large proteins or protein assemblies are not amenable for analysis, or crystallization is unsuccessful, especially for membrane proteins. To date, the high-resolution techniques X-ray crystallography and NMR spectroscopy provide the most highly resolved protein three-dimensional structures. Many thousands of protein structures had been solved to atomic level so far. However, both techniques suffer from a number of limitations, as described above. Therefore, alternative methods for the structural characterization of proteins and protein complexes have been developed that compensate for these shortcomings.

During the past few years, mass spectrometry has been increasingly employed for structural studies of proteins and protein complexes. The combination with chemical cross-linking, a method of covalently joining near-neighbor functional groups of amino acid residues, represents a promising alternative to existing methods for structural analysis of proteins. From the distance restraints imposed by the cross-linking reagents, low-resolution models of the proteins under investigation can be created by computational methods.

1.1.1 Analytical Methods for Protein Three-Dimensional Structure Analysis

1.1.1.1 High-Resolution Techniques

X-ray Crystallography

In the late 1950s and early 1960s, protein tertiary structure elucidation at the atomic level was successfully accomplished for haemoglobin and myoglobin by X-ray crystallography, for which Max Perutz and John Kendrew received the Nobel Prize in Chemistry in 1962 [Kendrew & Perutz, 1957].

Obtaining single crystals is a prerequisite for successfully recording diffraction patterns caused by the diffraction of X-rays by the atoms within the crystal, from which the protein structure can be deduced. About 85.8 % of all protein structures deposited in the PDB was solved by X-ray crystallography [www.rcsb.org/pdb]. However, this method can be time-consuming and requires rather large protein quantities for crystallization, if crystallization is successful at all. Furthermore, the crystal structure of a protein does not necessarily correspond to its structure in solution. Nevertheless, to date X-ray crystallography represents the most powerful method for obtaining protein structures at high resolution.

Nuclear Magnetic Resonance Spectroscopy

Nuclear Magnetic Resonance (NMR) Spectroscopy of proteins was pioneered by Kurt Wüthrich [Wüthrich, 2003] and established as a technique for the structural characterization of proteins in solution at the atomic level. Wüthrich received the Nobel Prize in Chemistry in 2002, which he shared with Fenn and Tanaka "for the development of methods for identification and structure analyses of biological macromolecules" [http://nobelprize.org/]. NMR spectroscopy requires large sample amounts and data analysis is time-consuming due to the complexity of the recorded spectra. Typically, proton NMR or NMR of isotope-labeled proteins is performed, for which recombinantly expressed proteins enriched with ^{13}C or ^{15}N are required [Lottspeich & Zorbas, 1998].

About 13.7 % of PDB protein structures were solved by NMR spectroscopy [www.rcsb.org/pdb]. The size limitations for NMR spectroscopy are currently well below 70 kDa [Wüthrich, 2003], but methods are developed that allow for studying large proteins and protein complexes [Horst *et al.*, 2006].

1.1.1.2 Mass Spectrometry – Based Approaches

Mass spectrometry (MS) was pioneered by the work of J. J. Thomson [Thomson, 1906]. However, almost one century was to pass, before, with the development of the soft ionization techniques matrix-assisted laser desorption / ionization (MALDI) [Karas & Hillenkamp, 1988]

and electrospray ionization (ESI) [Fenn *et al.*, 1989] in the late 1980s, mass spectrometry evolved as analytical technique in protein science.

With recent technological advances in the MS field, mass spectrometry has become a complementary technique for protein structure determination.

Combining mass spectrometry with hydrogen / deuterium (H/D) exchange experiments has proven successful for gaining knowledge of protein topologies [Zhang & Smith, 1993] and protein-protein interaction surfaces [Hasan *et al.*, 2002], as well as of conformational changes [Engen & Smith, 2001]. Mass spectrometry can also be employed for the analysis of non-covalent protein complexes. This is accomplished by ESI-MS under mild conditions and allows for determination of dissociation constants and the stoichiometry of protein complexes [Loo, 1997]. Another method for obtaining structural information on proteins and protein complexes is surface topology mapping. Lysine residues were modified by aminoacetylation and arginine residues can be modified by 1,2-cyclohexanedione as exemplified for lysozyme [Suckau *et al.*, 1992]. MS analysis of enzymatically digested modified proteins allows for drawing conclusions on surface topologies from the extent of modification.

Chemical cross-linking experiments in combination with MS represent another promising approach for the structural characterization of proteins and protein complexes [Sinz, 2003 & 2006, *Back et al.*, 2003]. In principle, two different approaches of chemical cross-linking combined with MS analysis can be distinguished. In the so-called "bottom-up" approach the cross-linked protein(s) are enzymatically digested prior to analysis with MS [Trester-Zedlitz *et al.*, 2003, Dihazi & Sinz, 2003, Schulz *et al.*, 2004, Kalkhof *et al.*, 2005], whereas the "top-down" approach directly analyzes the intact cross-linked protein (s) employing mass spectrometric fragmentation techniques, such as collisionally-activated dissociation (CAD), electron capture dissociation (ECD), or infrared multiphoton dissociation (IRMPD) [Novak *et al.*, 2003 & 2005, Kruppa *et al.*, 2003].

Combining chemical cross-linking experiments with mass spectrometric analysis possess a number of decisive advantages: When analyzing the proteolytic digestion mixtures of the cross-linking reaction mixture in a "bottom-up" approach the size of the protein or protein complex under investigation is theoretically unlimited, as it is the peptides that are analyzed. The analysis is generally fast and requires only minute amounts of analyte.

1.1.1.3 Further Methods

Besides the high-resolution techniques and mass spectrometry, other methods exist that are employed for the elucidation of protein structures. As mentioned above electron cryo-microscopy [Baumeister & Steven, 2000] is gaining in importance and is able to provide protein structures at a resolution well below 10 Å [www.rcsb.org/pdb]).

Further methods employed for obtaining structural information on proteins are for example neutron scattering [Gabel *et al*., 2002], FT-IR spectroscopy [Hering *et al*., 2002], FRET [Barth *et al*., 1998], and circular dichroism [Barth *et al*., 1998, Baumruk *et al*., 1996].

1.1.2   Computational Protein-Protein Docking

Computational protein-protein docking describes the process of computing the three-dimensional structure of a protein complex based on its known constituent proteins. It is increasingly employed for providing three-dimensional structures of protein complexes [Back et al., 2003, van Dijk *et al*. , 2004]. Chemical cross-linking combined with MS [Taverner *et al*., 2002, Lacroix *et al*., 1997], H/D exchange experiments [Anand *et al*., 2003], and mutagenesis studies [Tung *et al.*, 2002] have readily been employed for providing the data for successfully deriving structural models by protein-protein docking.

However, one has to be aware that the restraints imposed between functional groups of near-neighbor amino acids by covalent bonding through chemical cross-linking are by no means as numerous as those obtained from NMR spectroscopy. This is why only low-resolution structures are obtained from chemical cross-linking experiments. Nevertheless, the distance restraints provide valuable insights into protein structure and function.

For protein-protein docking the structures of the constituent proteins have to be available. In principle, two different approaches can be distinguished: In the first, at least one of the two protein components used for docking is derived from an already existing protein complex structure. Another approach makes use of the structures of the free proteins. These two approaches are termed "bound" and "unbound" docking, respectively. Furthermore, it is necessary to decide whether the proteins are kept rigid or whether a certain extent of flexibility is allowed. Moreover, There are several sampling (e.g. Monte-Carlo [Knegtel *et al*., 1994, Abagyan *et al*., 1994, Morris *et al*., 1998] or molecular dynamics (MD) [Dominguez *et al*., 2003]) methods to chose from.

Quite a number of different programs allowing for protein-protein docking are available: Rosetta [Gray *et al., 2003*, Daily *et al*., 2005], Haddock [Dominguez *et al*., 2003], Autodock [Morris *et al.*, 1998], to name only a few.

1.2 Mass Spectrometry

1.2.1 Ionization Methods

1.2.1.1 *Matrix-Assisted Laser Desorption / Ionization (MALDI)*

Matrix-assisted laser desorption / ionization (MALDI) is a development of the direct laser desorption ionization (LDI) [Gross, 2004] of small organic molecules that was initially developed in the 1970s.

In the middle of the 1980s, Tamio Yoshida, co-worker of Koichi Tanaka, had the idea of adding cobalt ultra-fine metal powder to the sample for reducing molecular photodissociation of the sample induced by direct laser irradiation. Tanaka further developed this approach by adding glycerin [Tanaka *et al.*, 1988] and together they submitted a patent (JP01769145) on this preparation technique. Tanaka was awarded the Nobel Prize in Chemistry in 2002.

In the late 1980s, Karas and Hillenkamp [1988] demonstrated that adding a small molecular weight organic matrix to an analyte was highly favorable for the analysis of biological samples.

In MALDI, the analyte is first embedded in a layer of small organic molecules, called matrix, that exhibit a strong UV absorption at the laser's wavelength (mostly 337 nm). Preparation of the analyte / matrix mixture involves dissolving the analyte in a solvent / matrix mixture and subsequent deposition of the analyte / matrix / solvent mixture on a solid surface. This preparation technique is called *dried-droplet* technique. The most commonly employed matrix for peptide analysis is α-cyano-4-hydroxy cinnamic acid. Other matrices used for protein analysis are 2,5-dihydroxy benzoic acid (DHB) and 3,5-dimethoxy-4-hydroxy cinnamic acid (sinapinic acid). Following introduction of the target plate into the ion source of the mass spectrometer, short intense laser pulses are applied to the co-crystallized analyte / matrix deposit. The exact mechanism of the ionization process in MALDI is still under discussion [Karas & Krüger, 2000, Knochenmuss & Zenobi, 2003, Dreisewerd, 2003].

1.2.1.2 *Electrospray Ionization (ESI)*

In addition to MALDI, the other most important soft ionization technique suitable for transferring biological samples into the gas phase is electrospray ionization (ESI), which was established in the late 1980s by John B. Fenn [Fenn *et al.*, 1989], for which he was honored with the Nobel Prize in Chemistry in 2002.

Electrospray ionization is obtained from the application of a strong electric field to the analyte containing liquid, which passes through a capillary. A potential difference of 2-6 kV is applied between the capillary and the counter-electrode.

When a critical electric field strength is reached, instant formation of the Taylor cone at the end of the capillary and immediate ejection of a jet of liquid from the Taylor cone towards the

counter electrode is observed. From the liquid filament charged droplets are emitted. The evaporation of solvent and concomitant descent in droplet size leads to a charge increase on the surface of the droplet. When a certain threshold is reached (the so-called Rayleigh limit) and electrostatic repulsion exceeds the surface tension, small highly charged micro-droplets are ejected (Coulomb explosion). These droplets are again desolvated and emit even smaller droplets. The formation of desolvated ions from the nano-droplets is described by two different theories. In the *charged-residue model* described by Dole *et al*. [1968] ion formation results from solvent evaporation of nano-droplets containing a single ion. According to Iribarne and Thomson [1976] and their *ion evaporation model* ions are directly ejected from nano-droplets.

Recent studies [Gamero-Castano & Mora 2000, Mora 2000] provided evidence for both models so that probably an in-between process takes place. ESI typically generates multiply charged ions allowing for detection of higher molecular weight proteins and peptides below *m/z* 2000. The complexity of the mass spectrum due to the occurrence of multiply charged signals can be reduced by recalculation of the multiply charged to singly charged species by a process called deconvolution.

1.2.2   Mass Analyzers

1.2.2.1   *Time-of-Flight Mass Spectrometry*

In 1946, Stephens [Stephens, 1946] first described the time-of-flight (TOF) mass analyzer. About one decade later, the first commercial TOF mass spectrometer was available. However, it was not until the end of the 1980s when renewed interest in this kind of instrumentation has been sparked. The advent of MALDI as a soft ionization technique together with pulsed nature of LDI being perfectly suited for the TOF mass analyzer, formed the basis for this renewed interest. The ions, which are emitted from the MALDI source, are accelerated by a potential *U* and fly a defined distance *d* before reaching the detector. During the flight in the field-free region of the flight tube, ions of different mass-to-charge (*m/z*) values disperse. The *m/z* ratios are determined by measuring the time that the ions travel through the field-free region between source and detector. The flight is proportional to the square root of *m/z*.

Increased mass resolution can be achieved by applying a time delay between ion formation and extraction. Wiley and McLaren [Wiley & McLaren,1955] first introduced time-lag focussing in the 1950s, known today as delayed-extraction (DE). Delayed-extraction (10 to several hundred nanoseconds) compensates for the initial velocity distribution of the ion packet generated by MALDI, so that ions with identical *m/z* values arrive simultaneously at the detector.

Mass resolution can be further improved by employing an ion mirror, called reflector [Mamyrin, 1994]. The reflector consists of evenly spaced electrodes that retard the arriving ions.

Ions with identical *m/z* values, but different kinetic energies penetrate the reflector to different extent. Ions with higher kinetic energy penetrate deeper into the reflector, thus covering a longer flight path, resulting in simultaneous arrival of ions with same *m/z* values on the detector.

For conducting linear MALDI-TOFMS measurements of high-molecular-weight proteins the reflector voltage is switched off and the ions proceed to the detector behind the reflector.

MALDI-TOF mass spectrometry has been successfully used for the analysis of a wide range of different analyte molecules.

1.2.2.2 *Fourier Transform Ion Cyclotron Resonance Mass Spectrometry*

The fundamentals of modern Fourier transform ion cyclotron resonance mass spectrometry (FTICRMS) [Comisarow & Marshall, 1996] were laid in 1932 when Lawrence and Livingston [Lawrence and Livingston, 1932] showed that the angular frequency of the circular motion of ions moving perpendicular to a homogenous magnetic field is independent of the radii of the ions. In 1974 FTICRMS was introduced [Comisarow & Marshall, 1974] and since then continuous advances led the to superior resolution and mass accuracies seen today.

Under the influence of a strong magnetic field (up to 16 Tesla) the ions are forced on a circular path by the Lorentz force and the centrifugal force, affecting them in opposite directions. The dependency of the angular velocity ω_c of the cyclotron motion on the *m/z* value of the ions and the magnetic field B is described by the following *equation*: $\omega_c = B \cdot z/m$. Upon excitation of the circulating ions, the ion packages induce an image current each time they pass the detector plates at their distinct cyclotron frequency. The image current is recorded and Fourier transformation is employed for conversion of the transient into a frequency spectrum, which is in turn converted to a mass spectrum [Comisarow & Marshall, 1974].

1.2.3 MS Fragmentation Techniques

Tandem mass spectrometric analysis (MS/MS) provides sequence information on the protein or peptide under investigation by fragmentation. Different low-energy dissociation techniques are used for the fragmentation of the selected precursor ions. In collision-induced dissociation (CID), also referred to as collisionally activated dissociation (CAD), collisions between excited ions and a neutral gas results in fragmentation through conversion of the ion's kinetic energy into internal energy. CID produces b- and y-type ions [Biemann, 1990] (**Figure 1.1**). Other fragmentation techniques are (IRMPD) and electron-capture dissociation (ECD). ECD generates c- and z• -type ions.
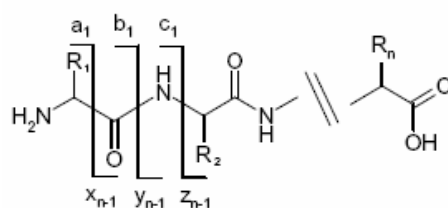
Figure 1.1: Nomenclature for Peptide Fragments.

MS/MS experiments are valuable in the analysis of cross-linked products, as they allow for an unambiguous assignment of the cross-linked product itself and for pinpointing cross-linked amino acid residue as could be shown for the CaM / adenylyl cyclase 8 peptide complex [Ihling *et al.*, 2006].

1.3 Chemical Cross-Linking

The process of chemically joining molecules by forming covalent bonds is called chemical cross-linking. Covalent linkages can be introduced between two ore more interacting molecules (intermolecular) or within one molecule (intramolecular) and provide valuable information on interaction surfaces, near-neighbor amino acid residues, protein conformation and surface topologies. Chemical cross-linking can be applied to various types of biological samples, such as proteins, peptides, nucleic acids, lipids, or cell surfaces.

Chemical cross-linking has a long established history in protein chemistry and a large number of reagents have readily been employed, many of which are commercially available [Hermanson, 1996, www.piercenet.com].

Most reagents consist of two reactive groups, which are interconnected by a spacer of variable length. The spacer defines the maximum distance between functional groups that are amenable to cross-linking. Hydrophobicity of the spacer influences both solubility of the reagent and its ability for penetrating into more hydrophobic regions of the protein or protein complex [Back *et al.*, 2003]. In addition to bifunctional reagents, trifunctional reagents, containing a third reactive moiety are available. Another group of cross-linking reagents comprises the so-called zero-length cross-linkers that do not introduce a spacer group, but directly connect functional groups of proteins [Hermanson, 1996].

When studying structural features of proteins or protein complexes it is crucial to maintain their native structure. The functional groups intended to be used for cross-linking have to be in close contact, i.e. within a distance of few Ångstroms in order to be cross-linked by a reagent with a defined spacer length. Another critical factor for cross-linking is that the functional groups to be cross-linked are accessible for the reagent.

Isotope-labeled cross-linkers greatly facilitate cross-linked product identification from the highly complex mass spectra obtained from enzymatically digested cross-linked proteins by searching for their distinct doublet isotopic pattern [Müller *et al.*, 2001, Schmidt *et al.*, 2005].

With the variety of available cross-linking reagents with respect to their reactivities and spacer lengths, combined with the advances in MS instrumentation, chemical cross-linking and mass spectrometry is gaining in importance as an alternative method for structural analysis of proteins and protein complexes [Sinz, 2006].

1.3.1 Bifunctional Cross-Linking Reagents
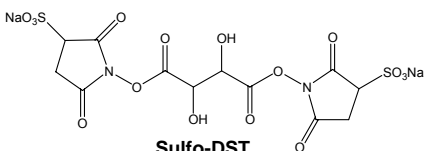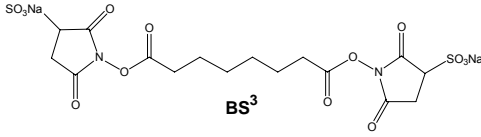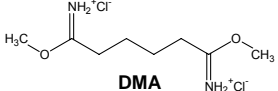
Bifunctional reagents contain two identical or two different reactive groups, which are separated by a spacer [Hermanson, 1996]. Homobifunctional cross-linking reagents contain two identical functional groups on either side of the molecule (**Table 1-1**). Heterobifunctional cross-linkers can be equipped with different combinations of reactive moieties (**Table 1-2**) and are largely employed in two-step cross-linking protocols.

In the following, different reactive groups of cross-linkers and their reaction mechanisms will be exemplified for the most commonly used reagents.

1.3.1.1 Amine-Reactive Cross-Linkers

Homobifunctional *N*-hydroxy sulfosuccinimide (sNHS) esters, such as disulfosuccinimidyl tartrate (sDST) and *bis*(sulfosuccinimidyl) suberate (BS³) (**Table 1-1**), are highly reactive towards α- and ε-amino groups, and their sulfonate groups provide sufficient water solubility. However, they are susceptible to hydrolysis.

Table 1-1: Amine-Reactive Homobifunctional Cross-Linkers.

Cross-Linking Reagent	Spacer Length	Name and Description
Sulfo-DST	6.4 Å	**Disulfosuccinimidyl tartrate** homobifunctional; reactive groups: sNHS esters (amine-reactive); cleavable by Na-periodate
BS³	11.4 Å	*Bis*(sulfosuccinimidyl) suberate homobifunctional reactive groups: sNHS esters (amine-reactive)
DMA	8.6 Å	**Dimethyl adipimidate** homobifunctional reactive groups: imidoesters (amine-reactive)

Upon cross-linking (**Figure 1.2**) sDST and BS³ produce amide bond cross-linked molecules, causing a mass shift of 113.995 u and 138.068 u, owing to insertion of the *spacer*,

respectively, whereas partially hydrolyzed cross-linkers exhibit a mass increase of 132.006 u and 156.079 u, respectively. SDST and BS[3] were used for the structural characterization of the calmodulin / melittin complex, as described in *chapter 2*. SDST was also employed for structural studies of the annexin A2 / p11 complex (chapter 4).
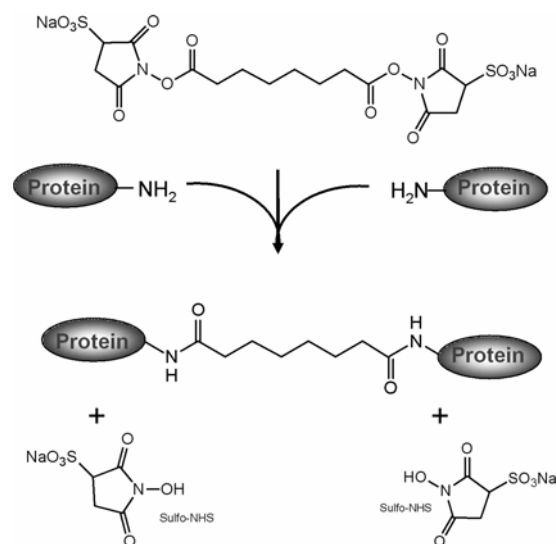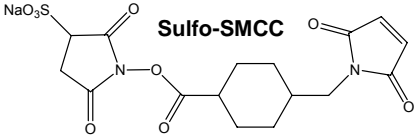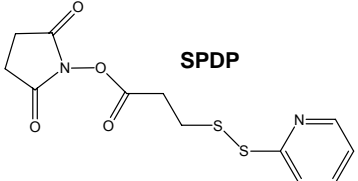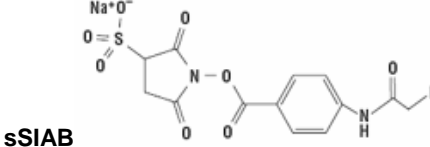
Figure 1.2: Cross-Linking Reaction Scheme of Amine-Reactive sNHS Esters. BS[3] is exemplarily presented.

Imidoester cross-linkers like dimethyl adipimidate (**Table 1-1**) react with primary amines to form amidine bonds. The resulting amidine is protonated, which is advantageous for subsequent mass spectrometric analyses and for sustaining an intact protein three-dimensional structure. Nonetheless, imidoester cross-linkers have been steadily superseded by NHS esters as the created amidine bonds are labile at high pH.

1.3.1.2 Sulfhydryl-Reactive Cross-Linkers

Sulfhydryl-specific cross-linkers are often employed as heterobifunctional amine- and sulfhydryl-reactive reagents (**Table 1-2**). The maleimide group of sulfo-SMCC (**Table 1-2**) reacts specifically with sulfhydryl groups at neutral pH and forms stable thioether bonds. Maleimides are also susceptible to hydrolysis. *N*-succinimidyl 3-(2-pyridylthio)propionate (**Table 1-2**) mediates thiol-disulfide exchange reactions through the reactive pyridyldisulfide. The reaction can be monitored via the absorption of released pyridine-2-thione at 343 nm. Another reactive group for sulfhydryl-specific cross-linking are haloacetyls. Most commonly iodoacetyl derivatives are employed, like for example in *N*-sulfosuccinimidyl [4-iodoacetyl]aminobenzoate (sSIAB) (**Table 1-2**). The reaction of the iodoacetyl group with a sulfhydryl occurs through nucleophilic substitution of iodine with a thiol, thereby resulting in a stable thioether bond. Iodoacetyl reactions should be performed in the dark to avoid generation of free iodine that could react with Tyr, His, and Trp residues.
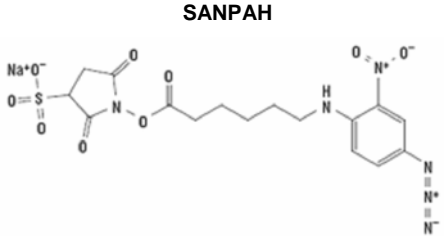
Table 1-2: Heterobifunctional Amine- and Sulhydryl-Reactive Cross-Linkers.

Cross-Linking Reagent	Spacer Length	Name and Description
Sulfo-SMCC	11.6 Å	**Sulfosuccinimidyl 4-(N-maleimidomethyl)-cyclohexane-1-carboxylate** heterobifunctional reactive group 1: sNHS ester (amine-reactive), reactive group 2: maleimide (sulfhydryl-reactive)
SPDP	6.8 Å	**N-Succinimidyl 3-(2-pyridylthio)propionate** heterobifunctional reactive group 1: NHS ester (amine-reactive), reactive group 2: pyridyldisulfide (sulfhydryl-reactive), mediates thiol-disulfide exchange reactions
sSIAB	10.6 Å	**N-Sulfosuccinimidyl[4-iodoacetyl]aminobenzoate** heterobifunctional reactive group 1: sNHS ester (amine-reactive), reactive group 2: iodoactyl (sulfhydryl-reactive)

1.3.1.3 Photo-Reactive Cross-Linkers

Photo-reactive cross-linkers react upon irradiation with long wave length UV light. The most commonly employed photo-reactive reagents are heterobifunctional cross-linkers with the photo-reactive site being either an aryl azide (**Table 1-3**), a diazirine, or a benzophenone [Weber & Beck-Sickinger, 1997]. Upon photolysis aryl azides form short-lived nitrenes that insert non-specifically into C-H, N-H, and double bonds. The reactions of aryl azides upon long wave length UV irradiation are schematized in **Figure 1.3**. Diazirines react via a reactive carbene created upon UV irradiation that inserts into C-H and heteroatom-H bonds. Benzophenone creates a biradical upon irradiation, which results in a newly formed C-C bond. Activation of benzophenone is reversible.

Table 1-3: Heterobifunctional Amine- and Photoreactive Cross-Linker

Cross-Linking Reagent	Spacer Length	Name and Description
SANPAH	18.2	**N-succinimidyl-6-[4'-azido-2'-nitrophenylamino]hexanoate** heterobifunctional reactive group 1: NHS ester (amine-reactive), reactive group 2: nitrophenylazide (photoreactive)

N-succinimidyl-6-[4'-azido-2'-nitrophenylamino]hexanoate (SANPAH, **Table 1-3**) possesses an amine-reactive NHS ester and a photo-reactive nitrophenyl azide as functional groups.

SANPAH was used in two-step cross-linking for the identification of binding partners of the annexin A2 / p11 complex (chapter 5). The two-step strategy involves tagging the protein for which the interaction partners are to be determined via the amine-reactive NHS ester in a first step. The such labeled protein is then added to a protein mixture and the binding partners are allowed to interact. At this point, photo-cross-linking is induced by irradiation with UV-light at 320-350 nm to covalently attach binding proteins.
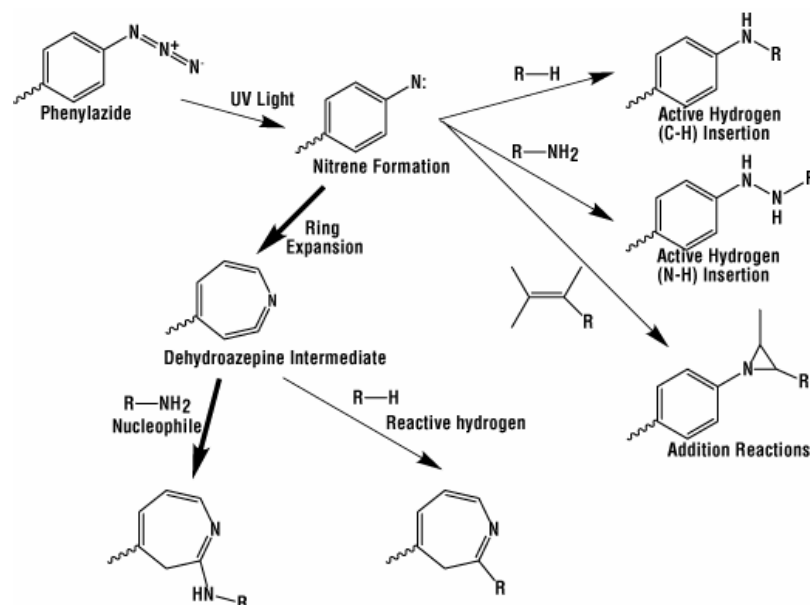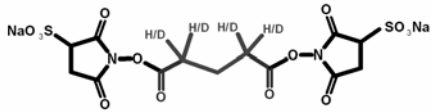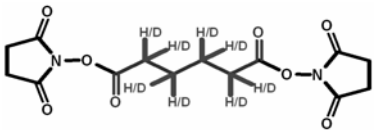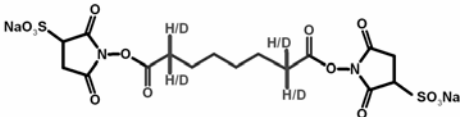
Figure 1.3: Reaction Scheme of Photoreactive Aryl Azides.
Scheme taken from www.piercenet.com

1.3.1.4 Isotope-Labeled Cross-Linkers

An important feature of cross-linking reagents is stable isotope-labeling in case the cross-linked products are analyzed by MS [Sinz, 2006b]. Hydrogen atoms of the spacer group are replaced by deuterium atoms and both non-deuterated and deuterated species are employed simultaneously in a fixed ratio. The signals of cross-linked products in the mass spectra are spaced at 4.025 (four-times deuterated) or 8.05 amu (eight-times deuterated), thus greatly facilitating cross-linked product identification by the characteristic doublet signals. Table 1-4 lists the three isotope-labeled reagents that were employed for structural studies of the annexin A2 / p11 complex (chapter 4). They are exclusively homobifunctional reagents with NHS esters targeting primary amines.

Table 1-4: Isotope-Labeled Cross-Linking Reagents

Cross-Linking Reagent	Spacer Length	Name and Description
BS^{2}G-d_0/d_4	7.7 Å	***Bis*(sulfosuccinimidyl) glutarate-d_0** **Bis(sulfosuccinimidyl) 2,2,4,4 glutarate-d_4** homobifunctional sNHS esters (amine-reactive), isotope-labeled
DSA-d_0/d_8	8.9 Å	**Disuccinimidyl adipate-d_0** **Disuccinimidyl adipate-d_8** homobifunctional NHS esters (amine-reactive), isotope-labeled
BS^3-d_0/d_4	11.4 Å	***Bis*(sulfosuccinimidyl) suberate-d_0** **Bis(sulfosuccinimidyl) 2,2,7,7 suberate-d_4** homobifunctional sNHS esters (amine-reactive), isotope-labeled

1.3.2   Trifunctional Cross-Linkers

In addition to the two reactive sites present in bifunctional cross-linking reagents, trifunctional reagents contain a third functional group. This third moiety typically constitutes a biotin group used for affinity purification of cross-linker-containing species. Less usual is the application of trifunctional reagents with three reactive groups as this leads to extensive aggregation of proteins and complicates data evaluation.

Sulfosuccinimidyl-2-[6-(biotinamido)-2-(p-azidobenzamido)hexanoamido]-ethyl-1,3‘-dithiopropionate (sulfo-SBED, **Figure 1.4**) is a representative of trifunctional reagents, containing an amine-reactive sNHS moiety, a photoreactive aryl azide, and a biotin group.
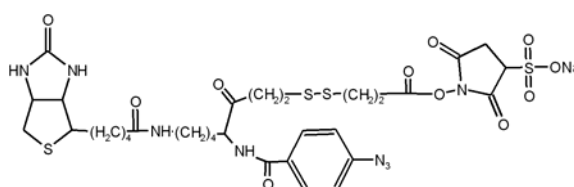
Figure 1.4: The Trifunctional Cross-Linking Reagent Sulfo-SBED.

Employing trifunctional reagents with an incorporated biotin moiety greatly facilitates correct cross-linked product assignment due to selective enrichment of cross-linked species from the complicated enzymatically digested reaction mixtures [Hurst *et al.*, 2005, Sinz *et al.*, 2005]. The biotin moiety can be released by cleaving the contained disulfide under reducing conditions. Other examples of trifunctional reagents containing a biotin group are 2-{N2-[Na-benzoyl benzoicamido-N6-6-biotinamidocaproyl]lysinylamido}ethyl-2’-(N-sulfosuccinimidylcarboxy)ethyl

disulfide and 5-[2-biotinylamidoethyl-dithiopropionamido]-3,7-diaza-4,6-diketanonanoic acid *bis-N*-sulfosuccinimidyl ester. The former contains an amine-reactive sNHS group and a photoreactive benzophenone group, while the latter one is equipped with two sNHS groups.

1.3.3   Zero-Length Cross-Linkers

In contrast to the above described cross-linking reagents that introduce a *spacer* between the cross-linked amino acid residues, zero-length cross-linkers do not introduce a spacer, but directly connect suitable near-neighbor functional groups. The most prominent representative of zero-length cross-linkers is 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC), which is commonly employed in combination with *N*-hydroxysulfosuccinimde (sNHS) (**Figure 1.5**).

EDC reacts with the carboxylic acid group of an aspartic or glutamic acid or the free C-terminus of a protein to form a short-lived active *O*-acylurea intermediate. Then, sNHS reacts with the EDC-active ester complex, extending the half-life of the activated carboxylate. In the presence of a lysine ε-amino group or the protein's free N-terminus an amide bond is formed, while sNHS is released, leading to the loss of an H_2O molecule and thus to a mass decrease of 18.011 u per cross-link.
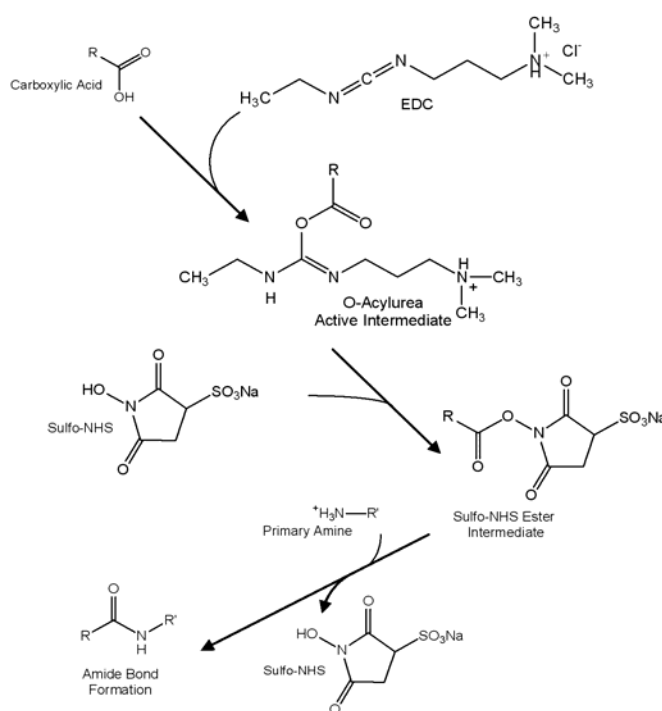
Figure 1.5: Reaction Scheme of the Zero-Length Cross-Linker EDC / sNHS. Scheme adapted from Hermanson, 1996.

EDC / sNHS were employed for chemical cross-linking of the calmodulin / melittin complex (chapter 2).
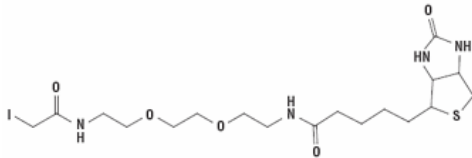
1.3.4 Protein Labeling Reagents

The same reactive principles underlying cross-linking reagents are present in protein labeling reagents, which possess only one reactive site.

One important modification reaction is the introduction of sulfhydryl groups into proteins through reaction with primary amines by 2-iminothiolane (Traut's reagent) [Hermanson, 1996].

The biotinylation of proteins via labeling reagents is a useful means for subsequent affinity purification via avidin. Biotinylation of the proteins with the sulfhydryl-specific reagent (+)-biotinyl-iodoacetamidyl-3,6-dioxaoctanediamine (PEO-iodoacetyl biotin) (**Table 1-5**) is highly feasible for two reasons: it can be employed in aqueous solutions, the polyethylene oxide chain confers water solubility, and due to it reacting with sulfhydryls, other functional groups in the protein remain unoccupied and thus are available for succeeding reactions.

Table 1-5: Sulfhydryl-Specific Biotinylation Reagent

Cross-Linking Reagent	Spacer Length	Name and Description
PEO-iodoacetyl biotin	24.7 Å	**(+)-Biotinyl-iodoacetamidyl-3,6-dioxaoctane diamine** reactive towards sulfhydryls biotin moiety allows affinity purification of labeled protein nucleophilic substitution of iodine with a thiol group (see chapter 1.3.1.2)

1.4 Analytical Strategy

The "bottom-up" analytical strategy for the structural characterization of proteins or protein complexes by chemical cross-linking combined with mass spectrometric analysis is depicted in **Figure 1.6**. In the first step, chemical cross-linking reagents are employed to covalently attach interacting amino acid residues within proteins or protein complexes via their reactive functional groups. The cross-linking reaction mixtures of the intact protein (intramolecular cross-linking) or protein complex (intermolecular cross-linking) are analyzed by MALDI-TOFMS for monitoring the extent of cross-linking. The reacted protein or protein complex is then separated by one-dimensional gel electrophoresis under denaturing conditions (SDS-PAGE). At this step the protein complex subunits that are not covalently cross-linked fall apart and are thus separated from cross-linked species. The gel bands of interest are excised and subjected to *in-gel* enzymatic proteolysis, which yields highly complicated peptide mixtures. From intramolecular cross-linking intrapeptide cross-linked products as well as interpeptide intramolecular cross-linked products are observed. In addition to these, intermolecular cross-linked products between

two proteins are created. Moreover, peptides, which are modified by a partially hydrolyzed cross-linker are observed. From mass spectrometric analysis (LC/ESI-MS and LC/ESI-MS/MS) of the proteolytic peptide mixtures peaklists are generated and comparison of the obtained *m/z* values to theoretical peptides and cross-linked products allows assignment of cross-linked products. The sequence information on cross-linked products provided by MS/MS data facilitates cross-linked products and enables identification of the attached amino acid residues. Thus, the sites of interaction are revealed and one can estimate the approximate distance of the amino acid residues involved in the interaction.
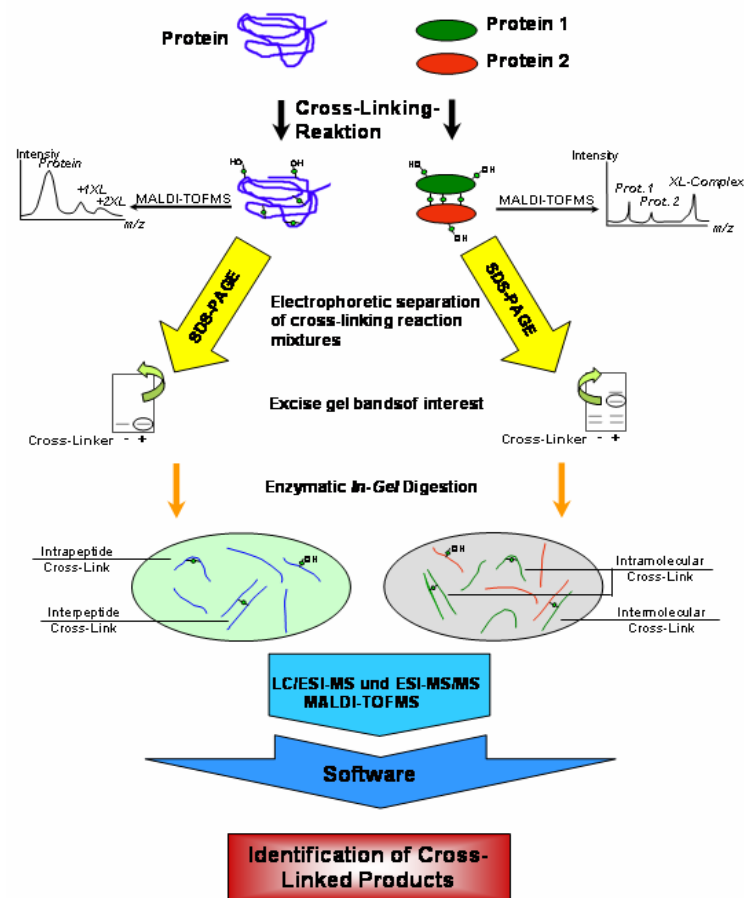
Figure 1.6: Bottom-Up Strategy for Analyzing Cross-Linked Products.

1.5 Studied Protein Complexes

Chemical cross-linking and mass spectrometry were applied for the structural characterization of two different protein complexes:

- the ~ 20 kDa complex between the Ca^{2+}-binding protein calmodulin and its target peptide melittin, and

- the 100 kDa heterotetramer between annexin A2 and p11 (S100A10) that binds calcium-dependently to phospholipid membranes.

1.5.1 Calcium-Binding Proteins

Ca^{2+} acts as a ubiquitous signal molecule within the cell. In the normal state the intracellular concentration Ca^{2+} is very low (~ 100 nM). Ca^{2+} is stored within organelles at exceedingly higher concentrations, usually in the endoplasmic reticulum (sarcoplasmic reticulum in muscle cells), from where it is released upon a trigger signal. This spatial partition and the steep Ca^{2+}-gradient allow Ca^{2+} to exert its function as second messenger upon slight changes in Ca^{2+} concentration [Stryer, 1995].

1.5.1.1 EF-Hand Proteins

Many well-known Ca^{2+}-binding signaling proteins belong the family of EF-hand proteins, such as calmodulin, troponin C and S100 proteins [Carafoli, 2002], which typically undergo a calcium-dependent conformational change. The EF-hand is a helix-turn-helix structural motif in proteins that received its name from parvalbumin helices E and F flanking the Ca^{2+}-binding site [Kretsinger, 1987], thus assuming the shape of a grasping hand, with the forefinger and the thumb representing helices E and F, respectively. (**Figure 1.7 A**).
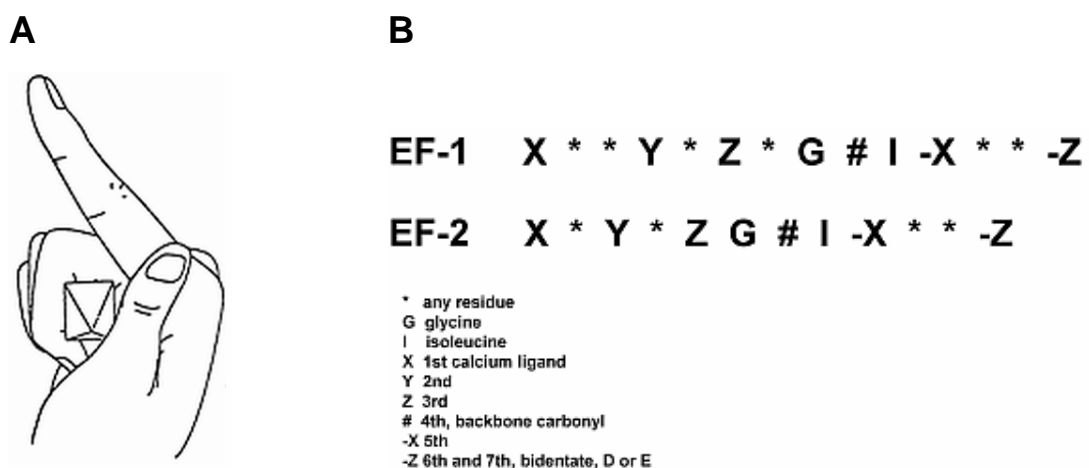
Figure 1.7: EF-hand motif. *A*, EF-hand (adapted from Kretsinger, 1973). *B*, Consensus sequences of EF-hand motifs.

The two alpha-helices are positioned roughly perpendicular to one another and are linked by a short loop region (usually about 12 amino acids) (**Figure 1.7 B**). The calcium ion is coordinated in a pentagonal bipyramidal configuration. The calcium ion is bound by oxygen atoms of five loop residues as well as of the peptide bond carbonyl oxygen. At position –Z an

invariant Glu or Asp provides two oxygens for Ca^{2+} coordination (bidentate ligand) [http://structbio.vanderbilt.edu/cabp_database].

1.5.1.2 Endonexin-fold

The Ca^{2+}-binding sites are termed type II and type III Ca^{2+}-binding sites and they con a 17-amino acid consensus sequence that is referred to as endonexin fold [Delmer *et al*., 1997]. The annexins consist of four (or eight, in the case of annexin A6) repeated segments of 70 amino acids, each forming a compact domain consisting of 5 α-helices [Liemannn & Huber, 1997]. These harbor the highly conserved 17-amino acid sequence comprising the endonexin fold with the consensus sequence GXGTDE.

1.5.2   Calmodulin / Melittin Complex

1.5.2.1   Calmodulin

Calmodulin (CaM) is a small (148 amino acids) acidic protein belonging to the class of EF-hand proteins, which is found ubiquitously in animals, plants, fungi, and protozoa [Crivici and Ikura, 1995]. It is highly conserved across all eukaryotes [Crivici and Ikura, 1995], emphasizing its exceptional role as a key component of the Ca^{2+} second-messenger system and thus being involved in controlling many metabolic pathways [Carafoli, 2002]. A rise in cytosolic Ca^{2+} concentration to ~ 500 nM results in activation of CaM, which then binds up to four Ca^{2+} ions. Overall, CaM adopts a dumbbell structure [Babu *et al.*, 1985 & 1988] consisting of two lobes, containing the EF-hand motifs, connected by a flexible central helix [Persecchini *et al.*, 1988, Barbato *et al.*, 1992] (**Figure 1.8 B**).

A   **B**
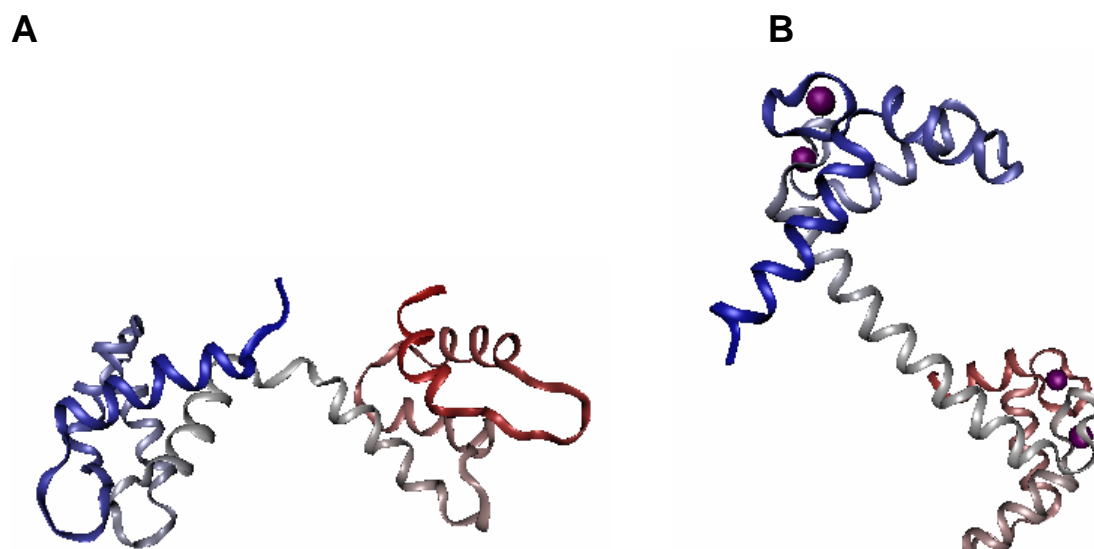
Figure 1.8: Calmodulin Structures. *A*, Apo-CaM (PDB entry 1DMO). *B*, Fully Ca^{2+}-loaded CaM (PDB entry 1CLL).

The CaM EF-hands comprise amino acid sequences 20-31 (DKDGDGTITTKE), 56-67 (DADGNGTIDFPE), 93-104 (DKDGNGYISAAE), and 129-140 (DIDGDGQVNYEE) [Swiss-Prot entry P02593]. While in its Ca^{2+}-free state (apoCaM) CaM the hydrophobic regions of CaM are buried within the protein (**Figure 1.8 A**), the conformational change induced by Ca^{2+}-binding (**Figure 1.8 B**) exposes methionine residues, thus forming two hydrophobic patches, one within each lobe [Bähler & Rhoads, 2003]. The propensity of CaM for binding a variety of different target peptides is caused by the presence of these hydrophobic patches and the ability of the central helix to function as a flexible tether.

1.5.2.2 Melittin

CaM is known to bind with high affinity to various target proteins and peptides, with dissociation constants in the low nanomolar range [Comte *et al.*, 1983]. One target is the small, 26-residue peptide melittin, the principal component of the honeybee (*Apis mellifera*) venom. Concomitant with complex formation, melittin adopts an α-helical conformation [Maulet & Cox, 1983] consisting of two α-helical segments interrupted by proline, with a bent rod-shaped appearance [Eisenberg *et al.*, 1980, Terwilliger & Eisenberg, 1983]. Melittin has been widely used as a tool for studying protein-protein interactions, and is known to interact with a number of proteins, including CaM [Kataoka *et al.*, 1989], myosin light chains [Malencik & Anderson, 1988], and calsequestrin [He *et al.*, 1993]. Melittin activates protein kinase C, Ca^{2+}/CaM-dependent protein kinase II [Raynor *et al.*, 1991], phospholipase A_2 [Fletcher & Jiang, 1993], and inhibits P-type ion-motif ATPases [Cuppoletti *et al.*, 1989, Cuppoletti, 1990, James *et al.*, 1988]. It has been demonstrated that melittin is one of the most potent inhibitors of CaM activity, and as such is also a potent inhibitor of cell growth [Hait *et al.*, 1983, Lee and Hait, 1985]. It has also been shown that melittin inhibits the melanotropin receptor in M2R melanoma cell membranes [Gest and Salomon, 1987], and a recent report has revealed that bee venom inhibits carcinoma cell proliferation *in vitro* and tumor growth *in vivo* [Orsolic *et al.*, 2003].

1.5.2.3 Calmodulin / Target Peptide Complexes

A common feature of most high-affinity CaM-binding targets is the presence of an amino acid sequence, which is capable of adopting the structure of an amphiphilic α-helix [Cox *et al.*, 1985]. With respect to CaM, target binding induces a conformational transition from the dumbbell structure to an overall globular-shaped complex with the two lobes of CaM coming in close proximity to one another [Clore *et al.*, 1993, O'Neil & DeGrado, 1990]. Thus, the target peptide is engulfed in a hydrophobic channel [Ikura *et al.*, 1992, Yuan *et al.*, 1998]. Hydrophobic interactions involve bulky hydrophobic anchor residues of the CaM-binding target and hydrophobic patches that are located within the two lobes of CaM [Ikura *et al.*, 1992, Vetter & Leclerc, 2003, Itakura & Iio, 1992]. A cluster of basic amino acid residues in the sequences of

CaM-binding peptides can also interact with the highly abundant negatively charged residues of CaM [Osawa *et al.*, 1999, Meador *et al.*, 1992] (**Figure 1.9**).

The α-helical peptide is clamped by the N- and C-terminal domains of CaM, each of which contains a hydrophobic cleft that can easily accommodate bulky hydrophobic anchor residues. Two channel outlets have been defined for CaM with channel outlet 2 (CO-2) being more negatively charged than channel outlet 1 (CO-1). Thus, CO-2 is thought to contribute to peptide binding orientation through electrostatic interactions with the peptide's basic cluster at its one side [Osawa *et al.*, 1999]. The two domains of CaM are shifted relative to each other with the N-terminal domain being more oriented towards CO-1 and the central linker, whereas the C-terminal domain of CaM is shifted towards CO-2. Therefore, CO-1 is partially covered by CaM's central linker causing CO-1 to be slightly smaller than CO-2.

Figure 1.9 depicts several target peptide sequences, as well as their respective orientations, within CaM / peptide complexes. In most previously known CaM / target peptide complexes, such as CaM complexed with skMLCK peptide [Ikura *et al.*, 1992], smMLCK [Meador *et al.,* 1992], and CaMKII [Meador *et al.,* 1993], the peptide α-helix is positioned in a way that the N-terminal portion of the peptide interacts with the C-terminal domain of CaM, whereas the C-terminal portion binds to the N-terminal domain of CaM (orientation B, **Figure 1.9**).

Peptide	Sequence	Spacing	
Orientation A			
Melittin	GIGAVLKVLTTGLPALISWIKRKRQQ	12	
rCaMKKα	VKLIPSWTTVILVKSMLRKRSFGNPF	14	
cCaMKKα	VRVIPRLDTLILVKAMGHRKRFGNPFR	14	
Orientation B			
Melittin	QQRKRKIWSILAPLGTTLVKLVAGIG	12	
skMLCK	LAGSSSIKKFRNAASVAIFNKKWRRK	12	
smMLCK	SSLRGIARVAHGTKQWKRRA	12	
CaMKII	SFNRTALMTTLIAGKLKRRANFKKL	8	
CaMKI	KRMHRVVATANFAQKWKSKAFNKKI	8	
C20W	KIQTQIRNLGRFWLIQGRRL	--	
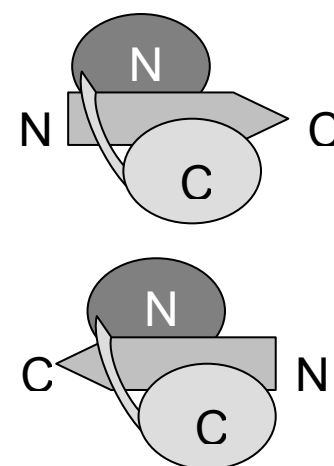
Figure 1.9: CaM-Target Peptide Orientations. Basic amino acid residues are highlighted gray and printed white, hydrophobic amino acid residues are highlighted gray and printed black. Spacing between hydrophobic anchor residues is given. Orientation scheme adapted from Osawa, 1999.

But there are also complexes, in which the peptide's orientation is reversed, as is the case for CaM / CaMKKα complexes [Osawa *et al.*, 1999, Kurokawa *et al.*, 2001]. For these peptides, the N-terminal portion interacts with the N-domain of CaM, whereas the C-terminal portion of the peptide binds to the C-terminal domain of CaM as shown in orientation A (**Figure 1.9**).

In addition to electrostatic interactions there appear to be steric determinants that underlie the orientation of the peptide, which is defined by the distance that separates the two

hydrophobic anchor residues in the target peptide. In melittin, as well as in the skeletal and smooth muscle MLCK peptides, there is a spacing of 12 residues between the two anchoring amino acids W and L (for melittin and smMLCK peptide) or W and F (for skMLCK peptide) (**Figure 1.9**). In rCaMKKα, however, two hydrophobic anchor residues are separated by a stretch of 14 amino acids, which causes the peptide to contain an additional well-defined β-hairpin-like loop that folds back into the complex [Osawa 1999]. The synthetic peptide C20W binds to only the C-terminal domain of CaM owing to the absence of a second hydrophobic anchor residue.

1.5.3 Annexin A2 / p11 Complex

1.5.3.1 Annexins

Annexins are a family of structurally related calcium- and phospholipid-binding proteins that are present in numerous eukaryotic species, including animals, plants, fungi, and protista, and reversibly bind to various cell membrane components [Morgan & Fernandez, 1997, Raynal & Pollard, 1994, Gerke & Moss, 2002, Moss & Morgan, 2004]. They mediate essential cellular processes via the ubiquitous second messenger calcium, however, it has been shown that membrane binding is not necessarily Ca^{2+}-dependent [Jost *et al.*, 1997]. Amongst the diverse biological functions of annexins are membrane trafficking, mitotic signaling, cytoskeleton rearrangement, DNA replication, and ion channel activity. Annexins are involved in a number of diseases, such as promyelotic leukemia, cardiovascular diseases, and cancer [Gerke & Moss, 2002, Gerke *et al.*, 2005, Rand, 2000]. Proteins of the annexin family contain a highly conserved core region with several homologous repeats and a flexible N-terminal domain that is not conserved. The N-termini of the different annexins vary both in length and amino acid composition and account for the functional diversity of annexins [Liemann & Huber, 1997].

Annexins are classified depending on their occurence: **A** vertebrate annexins, **B** non-vertebrate metazoans, **C** mycetozoa and fungi, **D** plants, and **E** alpha giardins [Moss & Morgan, 2004].

1.5.3.2 Annexin A2

Annexin A2 (ANXA2) is abundantly expressed in various tissues such as brain, spleen, kidney, lung, placenta, intestine, and the adrenal glands [Gerke & Moss, 2002]. Moreover, it is expressed in monocytes and macrophages [Brownstein *et al.*, 2004] and is involved in the assembly of plasminogen activators on cell surfaces [Hajjar & Krishnan, 1999], the regulation of vesicular traffic, endosome fusion, insulin signal transduction [Biener *et al.*, 1996], and RNA binding [Filipenko *et al.*, 2004]. Like other annexins, ANXA2 has a slightly curved, disc-like shape with a convex and a concave side (**Figure 1.10**).
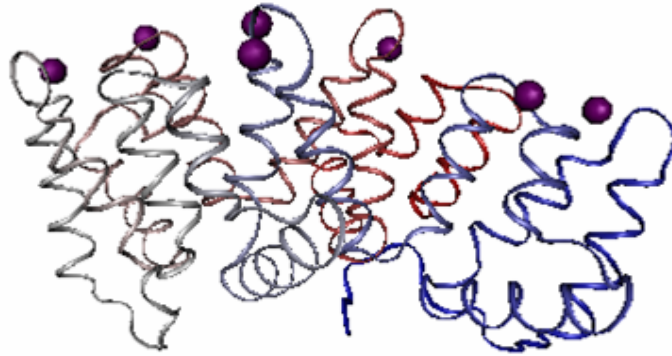
Figure 1.10: ANXA2. PDB entry 1W7B.

Calcium-binding and membrane attachment occur at the convex side of the core domain, whereas the acetylated α-helical N-terminus that mediates interaction with p11 is located at the opposite concave side [Gerke & Moss, 2002, Johnsson *et al.*, 1988, Becker *et al.*, 1990]. ANXA2 consists of a four-fold repeated segment of ca. 70 amino acids, the annexin repeat, containing a consensus sequence termed "endonexin fold", which harbors the Ca^{2+}-binding site [Liemann & Huber, 1997, Lewit-Bentley *et al.*, 2000]. ANXA2 is almost entirely α-helical and each domain consists of five helices (A-E), with four of the helices being oriented (anti-) parally and the fifth perpendicular to them [Liemann & Huber, 1997, Lewit-Bentley *et al.*, 2000]. ANXA2 forms a high-affinity heterotetrameric complex (A2t) with S100A10, which is also referred to as p11.

1.5.3.3   S100 Proteins

S100 proteins are low molecular weight proteins (9-13 kDa) that represent the largest subgroup of Ca^{2+}-binding proteins of the EF-hand type. S100 proteins are structurally related to CaM, however, they are, in contrast to the ubiquitous Ca^{2+} sensor CaM, expressed in a cell-type and cell cycle-specific manner [Donato, 2001]. Proteins of the S100 family have been implicated in a variety of intracellular processes, such as cell growth and motility, cell cycle regulation, transcription and differentiation. Extracellularly, they exhibit neurotrophic and antimicrobial properties (Donato). The name S100 proteins is derived from the fact that these proteins are 100% soluble in ammonium sulfate at neutral pH.

S100 proteins are involved in a number of diseases, including neurodegenerative disorders and cancer. They are used as markers for certain tumors and epidermal differentiation and hold potential as drug targets [Heizmann *et al.*, 1999].

Most S100 proteins form non-covalent dimer with two-fold axis symmetry. Ca^{2+}-dependent dimerization induces the exposure of the interaction surfaces on opposite sides of the dimer [Donato, 2001].

1.5.3.4   P11 (S100A10)

The S100A10 protein belongs to the S100 protein family and possesses a molecular weight of around 11 kDa. Therefore, it is also referred to as p11. Other synonyms for S100A10 are calpactin-1 light chain and annexin light chain [Swiss-Prot entry P04163 for *Sus scrofa*]. P11 is clearly distinct from all other S100 proteins as it does not undergo calcium-dependent conformational changes. Even in the absence of calcium, p11 is in a permanent "Ca^{2+}-on" state and, like most other S100 proteins, occurs as an anti-parallely packed non-covalent homodimer [Réty *et al.*, 1999]. P11 possesses anti-inflammatory properties through its interaction with cytosolic phospholipase A$_2$ [Wu *et al.*, 1997] and mediates translocation of ANXA2 to the cell surface [Deora *et al.*, 2004]. Furthermore, p11 targets ANXA2 to the cortical cytoskeleton of fibroblasts and adrenal chromaffin cells, enhances the ANXA2 dependent bundling of F-actin, and is thought to inhibit the extrinsic pathway of blood coagulation [Donato, 2001].

1.5.3.5   Annexin A2 / P11 Heterotetramer (A2t)

ANXA2 forms a heterotetrameric complex (A2t) with p11. The formation of the heterotetramer results from the attachment of two ANXA2 subunits to the p11 dimer [Liemann & Huber, 1997]. Within the cell, A2t is located on endosomal membranes and the plasma membrane where it exerts its role in endo- as well as exocytotic processes. A2t has also been shown to be involved in tight junction assembly [Lee *et al.*, 2004] and the perinuclear localization of *c-myc* mRNA [Mickleburgh *et al.*, 2005].

To date, no high-resolution structure of complete A2t exists. Structural data are available for the ANXA2 monomer with the first 19 N-terminal amino acids missing [Burger *et al.*, 1996, Rosengarth & Luecke, 2004] and the p11 dimer as well as for the complex between dimeric p11 and two synthetic N-terminal fragments (aa 1-11) of ANXA2 [Réty *et al.*, 1999]. Furthermore, a structural model of the ANXA2 / p11 complex has been created based on computational modeling using the structures of the N-terminally truncated ANX A2 monomer and the p11 / ANXA2 N-terminal peptide complex [Sopkova-deOliveira Santos *et al.*, 2000]. There are a number of reports on the characterization of the spatial organization and membrane binding of A2t, including site-directed mutagenesis [Kube *et al.*, 1992], cryo-electron microscopy [Lambert *et al.*, 1997], and scanning force microscopy [Menke *et al.*, 2004]. It has been frequently reasoned that ANXA2 and p11 not only occur as heterotetramer, but eventually form a heterooctameric complex via the interaction of two p11 dimers [Waisman, 1995, Lewit-Bentley *et al.*, 2000, Menke *et al.*, 2004].

2 Structure Determination of the Calmodulin / Melittin Complex

Employing chemical cross-linking in combination with high-resolution Fourier transform ion cyclotron resonance (FTICR) mass spectrometry, the interacting regions within the calmodulin / melittin (CaM / Mel) complex and thus the orientation of bound melittin were determined [Schulz *et al.*, 2004]. From the ambiguous distance restraints defined by the observed cross-linked products, low-resolution three-dimensional structure models of the CaM / Mel complex were generated by conjoined rigid body / torsion angle simulated annealing. The data provided evidence that CaM can recognize target peptides in two opposing orientations simultaneously.

2.1 Characterization of CaM and Melittin

The basis underlying the investigation of intermolecular interaction sites between CaM and melittin in the CaM / Mel complex is the assessment of detailed descriptions of their respective primary structures. The proteases trypsin and AspN as well as mixtures of LysC / AspN and trypsin / AspN were employed for enzymatic digestion of CaM and melittin and the digests were analyzed by nano-HPLC / nano-ESI-FTICR mass spectrometry. Full sequence coverages (100%) of both CaM and Mel were obtained from peptide mass fingerprint (PMF) analysis. The amino acid sequences of CaM and melittin are shown in **Figure 2.1**.

Calmodulin [a]:

*Ac*ADQLTEEQIA EFKEAFSLFD KDGDGTITTK ELGTVMRSLG QNPTEAELQD[50]
MINEVDADGN GTIDFPEFLT MMARKMKDTD SEEEIREAFR VFDKDGNGYI[100]
SAAELRHVMT NLGEKLTDEE VDEMIREADI DGDGQVNYEE FVQMMTAK[148]

Melittin [b]:

GIGAVLKVLT TGLPALISWI KRKRQQ-NH$_2$

Figure 2.1: Amino Acid Sequences of Calmodulin and Melittin. [a] Calmodulin was found to be N-acetylated (*Ac*-Ala) and to contain a trimethylated lysine in position 115 (underlined). [b] The C-terminus of melittin is amidated.

CaM was confirmed to be N-terminally acetylated and to contain a trimethylated lysine at position 115 [Watterson *et al.*, 1980]; the C-terminus of melittin was confirmed to be amidated.

Accordingly, exact mass measurements, after spectral deconvolution, of the intact protein and peptide performed by ESI-FTICRMS yielded a base peak at *m/z* 16790.921 for CaM (most abundant mass from simulation at *m/z* 16790.884, Δm = 0.037 u, 2.1 ppm) and a monoisotopic signal at *m/z* 2845.756 for melittin (calculated mass: [M+H]$^+_{monoisotopic}$ at *m/z* 2845.762, Δm = 0.006 u, 2.1 ppm) (data not shown).

2.2 Cross-Linking Reactions

For chemical cross-linking experiments between CaM and melittin the zero-length cross-linker 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC) in combination with *N*-hydroxysulfosuccinimide (sNHS), as well as the homobifunctional, amine-reactive cross-linking reagents disulfosuccinimidyl tartrate (sDST) and *bis*(sulfosuccinimidyl) suberate (BS^3) were used. Upon cross-linking sDST and BS^3 produce amide bond cross-linked molecules causing a mass shift of 113.995 and 138.068 u, respectively, whereas partially hydrolyzed cross-linkers exhibit a mass increase of 132.006 and 156.079 u, respectively. The reaction involving EDC and sNHS leads to the formation of an amide bond between a carboxylic acid group and a primary amine, resulting in the loss of an H_2O molecule and thus to a mass decrease of 18.011 u per cross-link. EDC and sNHS were employed at ratios of 1:1, 2:1, and 4:1, constituting molar excesses of 500x EDC / 500x sNHS, 1000x EDC / 500x sNHS, and 2000x EDC / 500x sNHS, respectively. BS^3 and sDST were applied at 50-fold molar excess over the protein concentration. In all cases, reaction times were 5, 15, 30, 60, and 120 minutes.

2.2.1 Gel Electrophoretic Separation of Cross-Linking Reaction Mixtures

The cross-linking reaction mixtures were separated by one-dimensional SDS-PAGE and the gels were stained with Coomassie Brilliant Blue. **Figure 2.2** shows a gel of melittin, CaM, and cross-linking reaction mixtures of an equimolar mixture of both proteins employing EDC / sNHS at a 2:1 ratio (1000-fold excess of EDC) after incubation times of 5, 15, 30, 60, and 120 minutes. CaM and melittin possess molecular weights of 16.8 kDa and 2.8 kDa, respectively.
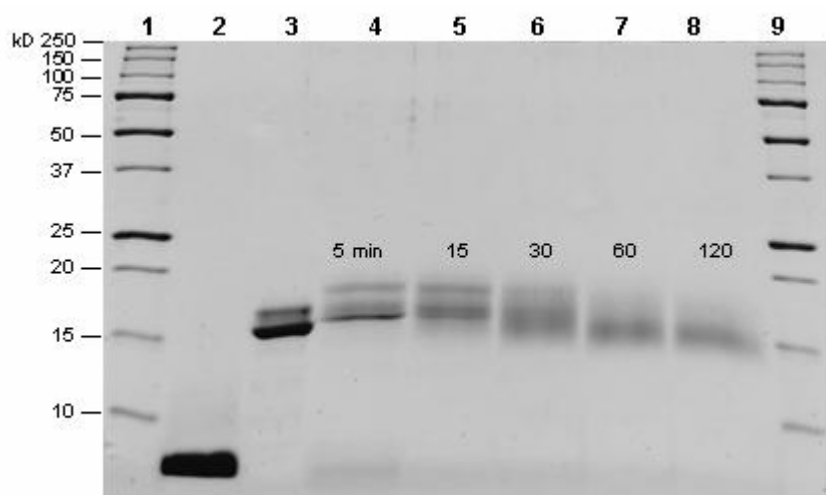
Figure 2.2: SDS-PAGE of the CaM / Mel Complex Cross-Linked with EDC / sNHS. 10µM CaM / 10µM melittin with EDC (10mM) / sNHS (5mM) (2:1). Lanes 1 and 9: molecular weight marker (kDa). lane 2: melittin (10µM); lane 3: CaM (10 µM), lane 4: 5min incubation time; lane 5: 15min; lane 6: 30min;lane 7: 60min; lane 8: 120min.

The cross-linked complex between CaM and melittin possess a molecular weight of ca. 19.6 kDa depending on the extent of chemical cross-linking. At an incubation time of 5 minutes, three distinct bands between 17 and 20 kDa are observed, as well as a faint band of melittin, indicating that cross-linking was not complete. After 15 minutes, only two bands can be clearly distinguished. When the incubation times are increased to 30, 60, and 120 minutes, the bands fuse and the mobility changes, indicating that alongside intermolecular cross-linking, intramolecular cross-linking had also occurred.

Similar findings were observed for the homobifunctional cross-linking reagent sDST. **Figure 2.3** depicts the gel electrophoretic separation of CaM, melittin, and the CaM / Mel complex cross-linked with 50-fold molar excess of sDST at reaction times of 5, 15, 30, 60, and 120 minutes.
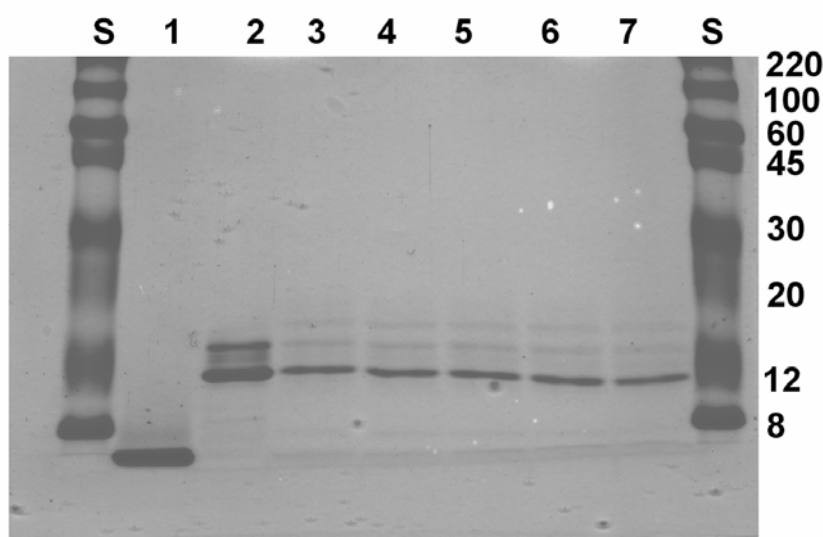
Figure 2.3: SDS-PAGE of the CaM / Mel Complex Cross-Linked with sDST. 10µM CaM / 10µM melittin, 50-fold excess sDST. S: molecular weight marker (kDa) (Colorburst). lane 1: melittin (10µM); lane 2: CaM (10 µM), lane 3: 5min incubation time; lane 4: 15min; lane 5: 30min;lane 6: 60min; lane 7: 120min.

After 5 minutes, formation of cross-linked complex is observed as indicated by a faint gel band migrating below ~ 20 kDa. In contrast to the reaction performed by EDC / sNHS, intramolecular cross-linking does not seems to occur at a significant extent. This results in very distinct gel bands of both the complex and monomeric CaM. Furthermore, it can be concluded that, in contrast to EDC / sNHS, chemical cross-linking efficiency with the homobifunctional NHS esters is not associated with incubation time.

The bands on the gels were excised separately whenever it was possible to tell them apart; otherwise the whole region was cut out, and the gel pieces were subsequently used for enzymatic *in-gel* digestion. Aggregation of proteins caused by excessive cross-linking was not observed for any of the cross-linking reagents as evidenced by the absence of any gel bands in the higher mass range (**Figures 2.2** and **2.3**).

2.2.2 Linear MALDI-TOFMS of Intact Cross-Linked Calmodulin / Melittin Complex

Linear MALDI-TOFMS using a Voyager-DE instrument was employed to estimate the extent of chemical cross-linking over the course of the cross-linking reaction. **Figure 2.4** shows MALDI-TOF mass spectra of the non-digested reaction mixtures after incubation times of 5, 15, 30, 60, and 120 minutes (bottom to top) with an EDC / sNHS ratio of 2:1 (1000-fold excess of EDC). At an incubation time of 5 minutes three distinct peaks appear at *m/z* 2848.4, 16792, and 19623, corresponding to the singly charged ions of melittin, CaM and the cross-linked CaM / Mel (1:1) complex, respectively. The observed average masses are in good agreement with the molecular weight of protonated melittin ([M+H]$^+_{average}$ 2847.5, Δm = 0.9 u), CaM with the post-translational modifications stated above ([M+H]$^+_{average}$ 16791.6, Δm $\approx$ 0.4 u), and the complex ([M+H]$^+_{average}$ 19638.1 (Δm $\approx$ 15 u), lacking 18 u due to one cross-link in the molecule).
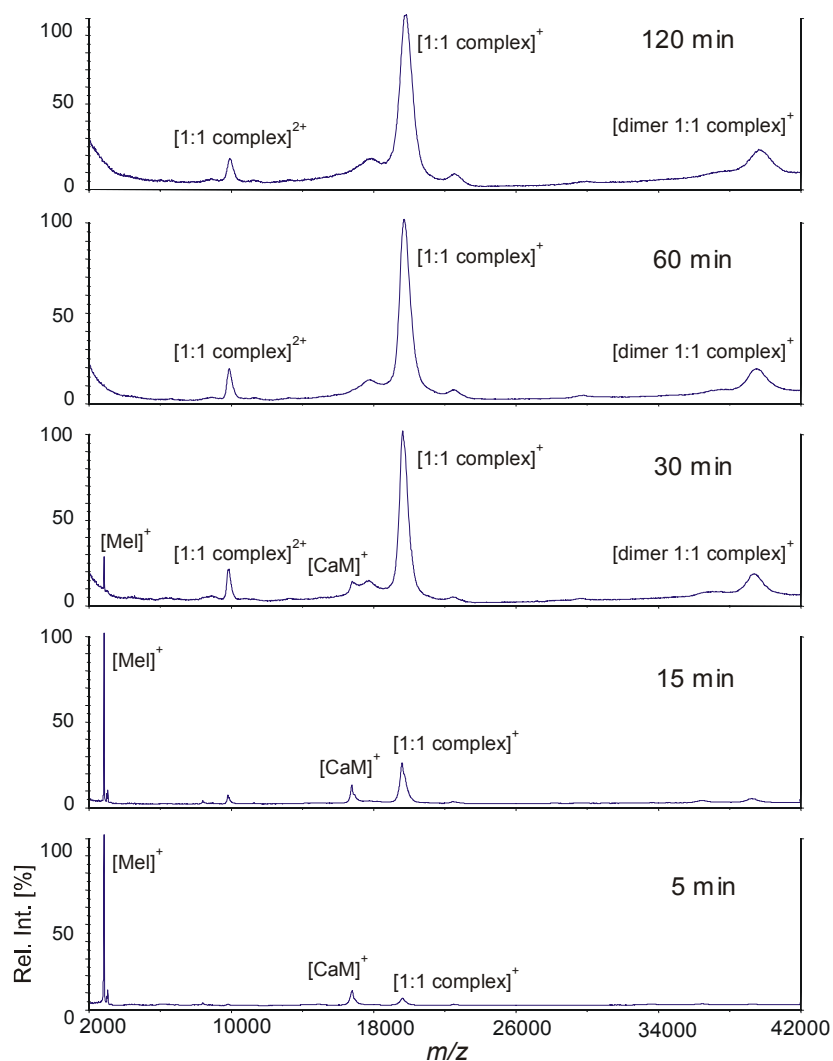
Figure 2.4: Linear MALDI-TOF Mass Spectra of the Intact Cross-Linked CaM / Mel Complex Cross-Linked with EDC /sNHS. MALDI-TOF mass spectra of the non-digested reaction mixtures after incubation times of 5, 15, 30, 60, and 120 minutes (bottom to top), 1000-fold excess of EDC, EDC / sNHS (2:1).

With increasing incubation time, the relative intensity of the singly charged signal of the CaM / Mel complex increases, whereas the relative intensity of the melittin signal steadily decreases and is no longer observed at incubation times beyond 60 minutes. Furthermore, the peak of the CaM / Mel (1:1) complex broadens indicating the formation of intramolecular cross-links in addition to intermolecular cross-linking products, when incubation times are increased. Average masses for the CaM / Mel (1:1) complexes range from *m/z* 19602 to 19833, steadily increasing during the time course. An explanation for a mass increase in the cross-linked complexes with increasing incubation time could be that the sNHS group remains attached to the protein if a reactive amino group is not present for the cross-linking reaction. Although MALDI preferentially generates singly charged ions, signals of the doubly charged ions of the CaM / Mel (1:1) complex are also observed between *m/z* 9780 to 9881. Signals of the singly charged ions of the complex dimer are visible between *m/z* 39239 to 39479.

Overall, the above described MALDI-TOFMS data are in agreement with the outcome of gel electrophoretic separation of the cross-linking reaction mixture (**Figure 2.2**). At incubation times of 5 and 15 minutes CaM can be clearly distinguished in both the gel and in the mass spectra, whereas after 30 minutes the cross-linked complex is the predominant species (**Figures 2.2** and **2.4**).

In contrast to these findings, the linear MALDI-TOF mass spectrum of the cross-linking reaction mixture employing 50-fold excess sDST at an incubation time of 30 minutes shows a different pattern of signals (**Figure 2.5**). Please note that the mass spectra for the other incubation times and for the cross-linking reagent BS[3] display identical signal patterns (data not shown). Being in agreement with the observation from gel electrophoretic separation (**Figure 2.3**), the MALDI-TOFMS measurements imply that the cross-linking yield of the homobifunctional reagents is much lower than obtained for EDC / sNHS cross-linking.
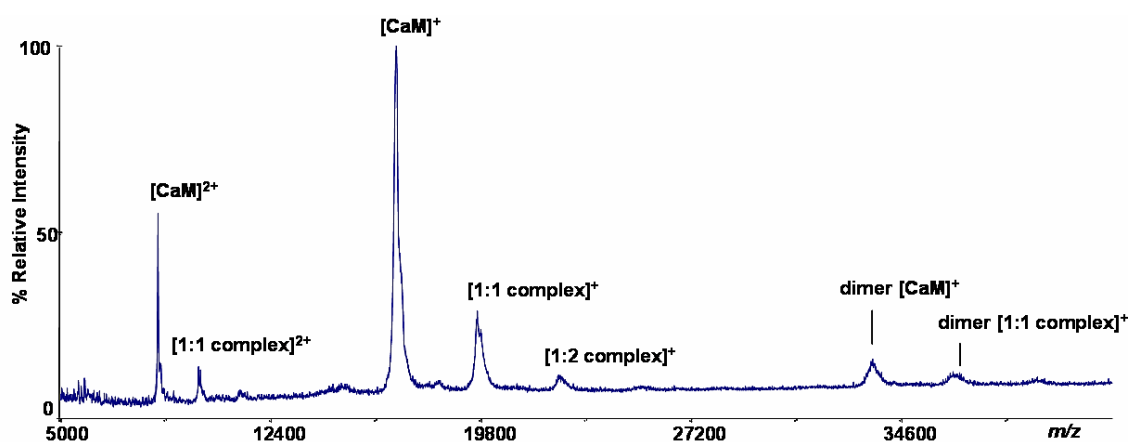
Figure 2.5: Linear MALDI-TOF Mass Spectrum of the Intact Cross-Linked CaM / Mel Complex Cross-Linked with sDST. 50-fold molar excess, 30 minutes reaction time.

The most abundant signal is that of singly charged monomeric CaM (**Figure 2.5**). The signal of the CaM / Mel (1:1) complex appears at *m/z* ~ 19800. A singly charged gas-phase adduct of the CaM / Mel (1:2) species was also observed. CaM and the CaM / Mel (1:1) complex dimers were additionally observed. Doubly charged signals of CaM and CaM / Mel species were observed.

2.2.3 *In-Gel* Enzymatic Digestion

Following separation of the cross-linking reaction mixtures by SDS-PAGE, the relevant gel bands were excised and the obtained gel pieces were subjected to enzymatic *in-gel* digestion with various proteases. For comparative purposes, gel bands of CaM and melittin were excised and treated in the same fashion as the bands of the cross-linked complex. For enzymatic *in-gel* digestion the proteases trypsin and AspN were applied in addition to combinations of trypsin/AspN and LysC/AspN. The latter two protease mixtures were exclusively applied to isolated CaM and melittin as well as to cross-linked products created using EDC in 1000-fold molar excess over the protein / peptide (EDC / sNHS ratio 2:1). Cross-linked complexes derived from cross-linking with sDST and BS[3], both applied in a 50-fold molar excess over the protein / peptide, were digested by trypsin alone.

2.3 Analysis of Cross-Linked Products

The peptide mixtures generated from enzymatic *in-gel* digestion were subsequently separated by nano-HPLC and analyzed by nano-ESI-FTICR mass spectrometry. These mixtures were rather complex, containing peptides from both CaM and melittin, as well as inter- and intra-molecular cross-linked products. The experimentally obtained monoisotopic masses were compared with calculated masses of peptides and cross-linked products employing the GPMAW and ASAP software packages. The theoretical cross-linked products for given protein sequences are calculated. Both software tools assign *m/z* values of peptides and cross-linked products generated from chemical cross-linking experiments. Information on cross-linking reagent, protease, and amino acid modifications has to be provided. One limitation of the GPMAW software is that variable amino acid modifications, such as methionine oxidation (MSO), cannot be automatically calculated. For this reason, several files of calmodulin sequences containing different combinations of MSO modifications had to be generated. This problem can be overcome by employing the ASAP software that is able to incorporate variable modifications. However, ASAP can only process cross-linking data for one single protein, but not for protein complexes.

2.3.1 Peptide Mass Mapping

To obtain additional information on the regions of CaM, which are involved in the interaction with melittin, peptide maps from cross-linked CaM / Mel (1:1) complexes were compared with those from non-cross-linked CaM and melittin. Sequence coverage maps were generated for samples that had been cross-linked with different amounts of EDC / sNHS, which had been *in-gel* digested with trypsin and endoproteinase AspN. With trypsin, full sequence coverage (100%) was obtained for CaM as well as for CaM / Mel complexes cross-linked with EDC / sNHS (1:1, 500-fold excess EDC / 500-fold excess sNHS and 2:1, 1000-fold excess EDC / 500-fold excess sNHS). However, when applying EDC / sNHS at a ratio of 4:1 (2000-fold excess EDC), the sequence coverage decreased to 89.2%. Only 132 out of the 148 amino acids of CaM were detected, with residues 75-90, which belong to the central α-helix of CaM, missing (data not shown).

Digestion with endoproteinase AspN led to different results as displayed in **Figure 2.6**.

```
I_not cross-linked
ADQLTEEQIA EFKEAFSLFD KDGDGTITTK ELGTVMRSLG QNPTEAELQD
MINEVDADGN GTIDFPEFLT MMARKMKDTD SEEEIREAFR VFDKDGNGYI
SAAELRHVMT NLGEKLTDEE VDEMIREADI DGDGQVNYEE FVQMMTAK
98.6% (146 of 148 amino acids covered)


II_500x EDC (1:1)
ADQLTEEQIA EFKEAFSLFD KDGDGTITTK ELGTVMRSLG QNPTEAELQD
MINEVDADGN GTIDFPEFLT MMARKMKDTD SEEEIREAFR VFDKDGNGYI
SAAELRHVMT NLGEKLTDEE VDEMIREADI DGDGQVNYEE FVQMMTAK
87.8% (130 of 148 amino acids covered)


III_1000x EDC (2:1)
ADQLTEEQIA EFKEAFSLFD KDGDGTITTK ELGTVMRSLG QNPTEAELQD
MINEVDADGN GTIDFPEFLT MMARKMKDTD SEEEIREAFR VFDKDGNGYI
SAAELRHVMT NLGEKLTDEE VDEMIREADI DGDGQVNYEE FVQMMTAK
45.3% (67 of 148 amino acids covered)


IV_2000x EDC (4:1)
ADQLTEEQIA EFKEAFSLFD KDGDGTITTK ELGTVMRSLG QNPTEAELQD
MINEVDADGN GTIDFPEFLT MMARKMKDTD SEEEIREAFR VFDKDGNGYI
SAAELRHVMT NLGEKLTDEE VDEMIREADI DGDGQVNYEE FVQMMTAK
16.9% (26 of 148 amino acids covered)
```

Figure 2.6: Sequence Coverages of Calmodulin from EDC / sNHS Cross-Linked CaM / Mel Complexes Digested with AspN. Different EDC / sNHS ratios were applied (1:1 to 4:1). Sequence coverages (given in %) decrease with increasing amounts of cross-linker.

Peptide mass mapping with AspN yielded a sequence coverage of 98.6% for CaM, lacking amino acids D56 and A57, because the mass of this peptide is too small to be detected. Covered sequences are displayed in red. For EDC / sNHS (1:1) sequence coverage of 87.8% was obtained with amino acids D56 and A57 as well as the sequence from residues 88 to 103 not being detected. With increasing amounts of EDC a decline in sequence coverage was observed. Digestion of the complex cross-linked with EDC / sNHS (2:1) yielded a sequence coverage of 45.3%, lacking the sequences comprising residues 20-49, 56-60, and 93-148. For the complex cross-linked with EDC / sNHS (4:1) only 16.9% of the CaM sequence was covered, lacking residues 20-86 and 93-148. These findings gave additional hints that the C- and the N-terminal domains as well as the central α-helix of CaM were involved in complex formation with melittin.

2.3.2   Cross-Linked Products Between Calmodulin and Melittin

As implied by the results from SDS-PAGE (**Figures 2.2** and **2.3**) and linear MALDI-TOFMS (**Figures 2.4** and **2.5**), the majority of cross-linked products between CaM and melittin were obtained from cross-linking with EDC / sNHS. The homobifunctional NHS esters sDST, and BS^3 also yielded peptides modified with partially hydrolyzed cross-linker.

2.3.2.1   Cross-Linked Products Obtained with EDC / sNHS

The results of the cross-linking experiments with EDC / sNHS are summarized in **Table 2-1**. From the reaction mixtures created with EDC / sNHS (1:1, 500-fold excess EDC), no cross-linked products were detected after digestion with trypsin; however, when employing an EDC / sNHS ratio of 2:1 (1000-fold excess EDC) (**Table 2-1**, samples III), a number of cross-linked products were observed. In the cross-linking reaction mixtures with EDC / sNHS (4:1) (**Table 2-1**, samples IV), three cross-linked products were found, one of which had not been observed when working with EDC / sNHS (2:1, 1000-fold excess EDC). In addition to using trypsin for digestion of the cross-linked complexes, endoproteinase AspN, trypsin/AspN, and LysC/AspN mixtures were used for digestion of the EDC / sNHS (2:1) cross-linking reaction mixtures. However, when using AspN as sole digestion enzyme, no cross-linked products were detected. Extending HPLC elution time of the digestion products with 95% acetonitrile to 60 minutes did not overcome this problem. Overall, most of the cross-linked products were observed in more than one sample (**Table 2-1**) and often the analogues containing an oxidized methionine (mass increase of 15.99 u) were detected as well, which further increased the certainty of a correct assignment.

Table 2-1: Cross-linked Products Between CaM and Melittin with the Cross-Linkers EDC / sNHS. EDC / sNHS were applied in 1000-fold/500-fold excess over the protein/peptide, (ratio 2:1, samples III) and 2000-fold/500-fold excess over protein (ratio 4:1, samples IV). Digestions were performed with the proteases trypsin (T), trypsin and AspN (T/A), LysC and AspN (L/A). Reaction times (in minutes) are given in numbers (5 to 120).

CaM	Modified Residue	Melittin	Modified Residue	[M+H]$^+_{calc}$	M+H]$^+_{exp}$	ppm	Sample	Protease(s)	
1-13	D2, E 6, E7, E11	8-22	K21	3212.757	3212.742 3212.750 3212.757 3212.759	5 2 0 1	III15 IV30 IV60 IV120	T	
14-30	E14, D20, D22, D24	8-22	K21	3493.894	3493.884	3	IV120	T	
78-86	D 78, D80, E82, E83, E84	1-7	G1	1731.876	1731.872 1731.872 1731.877	2 2 1	III5 III15 III30	T	
104-121	E114, D118, E119, E120	22-24	K23	2595.404	2595.393	4	III120	T/A	
107-126	E114, D118, D122, E119, E120, E123	23-24	K23	2701.365 2717.360	2701.363 2701.362 2717.356	1 1 1	III5 III15 III60	T	ORIENTATION A
118-130	E119, E120, D122, E123, E127, D129	22-24	K23	2003.981	2003.967 2003.970	7 6	III30 III120	T/A	
120-126	D 122, E123	23-24	K23	1191.615	1191.615	0	III120	T/A	
127-148	E127, D129, D131, D133, E139, 140	23-24	K23	2774.276 2790.271	2774.276 2774.268 2774.267 2790.271 2790.271	0 3 3 0 0	III5 III15 III30 III5 III15	T	
1-13	D2, E 6, E7, E11	1-7	G1	2202.165	2202.159 2202.157 2202.159 2202.166 2202.167	3 4 3 0 1	III5 III15 III30 IV60 IV120	T	ORIENTATION B
1-13	D2, E 6, E7, E11	1-7	G1	2202.165	2202.176	5	III30	L/A	
14-30	E14, D20, D22, D24	1-7	G1	2483.303	2483.298 2483.301 2483.307 2483.302	2 1 2 0	III5 III15 III30 III60	T	
127-148	E127, D129, D131, D133, E139, E140	1-7	G1	3144.486 3160.481	3144.478 3160.487	3 2	III15 III15	T	

In total, twelve cross-linked products were observed for the CaM / Mel (1:1) complex cross-linked with EDC / sNHS after digestion with trypsin, trypsin/AspN, and LysC/AspN (**Table 2-1**). As it is not possible in most cases to exactly identify the actually cross-linked amino acid, **Table 2-1** lists all the amino acids that are potentially involved in cross-linking.

Eight cross-linked products point conclusively to an orientation of melittin in the complex, which is inverted relative to that of the majority of the other CaM binding peptides studied to date (**Figure 1.9**). This binding mode was designated as orientation A (**Figure 1.9**). However, four additional cross-linking products between melittin and CaM were found, which did not match the orientation of melittin in the complex as suggested by orientation A (**Figure 1.9**). For example, residues 127-148 in the C-terminal domain of CaM (containing the acidic residues E127, D129, D131, D133, E139, and E140) were found to be cross-linked to residues 1-7 of melittin via the free N-terminus of the peptide corresponding to signals at *m/z* 3144.486 (M144 or M145 oxidized) and 3160.481 (both M144 and M145 oxidized) when using trypsin as digestion enzyme (**Table 2-1**). Conclusively, no cross-linked products were detected in the N-terminal domain and central linker region of CaM when employing the proteases trypsin and AspN simultaneously.

When using a LysC/AspN combination for digestion of the cross-linked complex, only one cross-linked product from samples of the CaM / Mel complex cross-linked with an EDC / sNHS ratio of 2:1 (1000-fold excess EDC) (**Table 2-1**) was detected. The identified cross-linked product comprised CaM residues 1-13 and melittin residues 1-7 (*m/z* 2202.165), pointing towards any of the four acidic residues in CaM 1-13 (D2, E6, E7, E11) being conjugated with the α-amino group of melittin's free N-terminus. The identical cross-linked product was observed in the tryptic digest. It is believed that the poor detection of cross-linked products, when using a LysC/AspN combination, is due to AspN not cleaving melittin, as melittin does not contain any acidic amino acids. Because the lysine residues (K7, K21, K23) of melittin are at least partly modified by the cross-linking reagent, LysC is also unable to cleave. Thus, the cross-linked products created in this manner are not amenable to observation in the optimum ESI-FTICRMS detection range between *m/z* 400 to 2000.

The combined results provide evidence that melittin is also positioned in a way, which corresponds to the peptide orientation found in the majority of the CaM / peptide complexes studied to date [26, 31-33, 35] and which is denoted as orientation B (**Figure 1.9**).

An example of the data quality is provided by **Figure 2.7**, which shows the deconvoluted nano-ESI-FTICR mass spectrum obtained from the tryptic digest of the CaM / Mel complex cross-linked with EDC / sNHS (2:1) (1000x EDC) after an incubation time of 15 minutes.
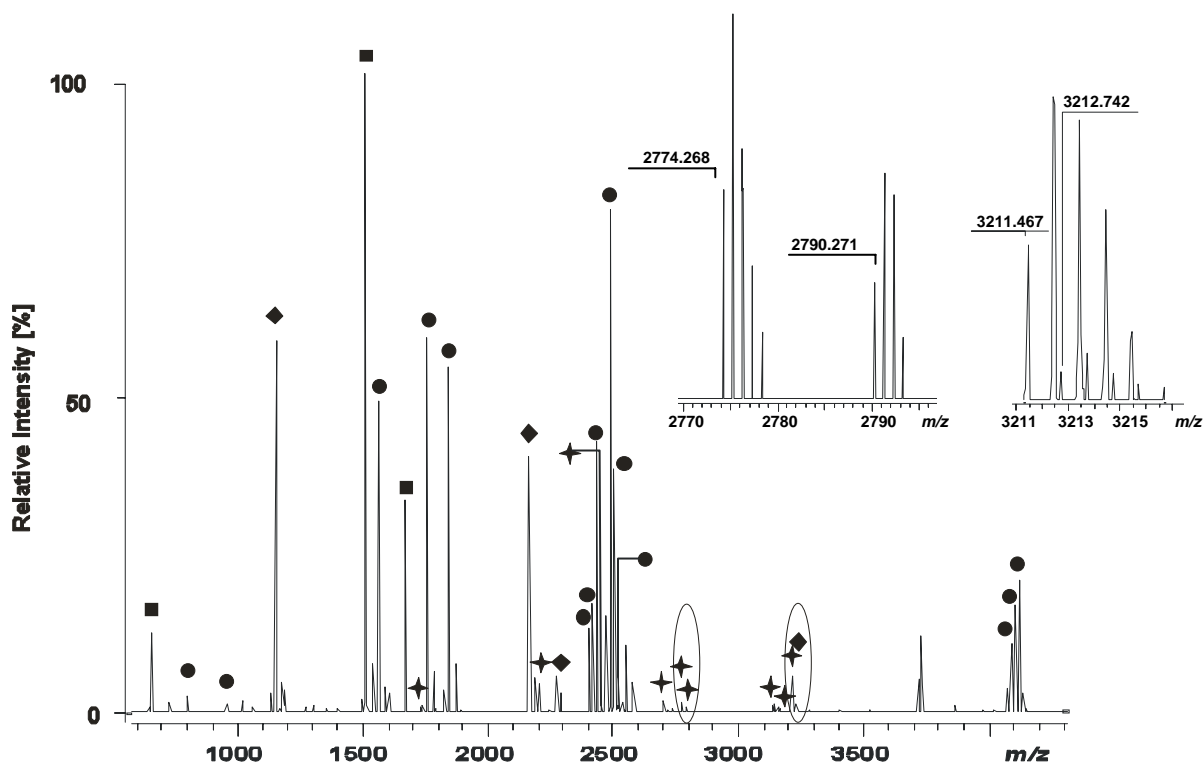
Figure 2.7: Nano-ESI-FTICR mass spectrum of the tryptic peptide mixture from the CaM / Mel complex cross-linked with EDC / sNHS (2:1) at an incubation time of 15 minutes. Star: cross-linked product, square: melittin peptide, circle: CaM peptide, diamond: autolytic tryptic peptide. The magnified inserts show a cross-linked product between CaM residues 1-13 and melittin residues 8-22 ([M+H]$^+_{calc}$ *m/z* 3212.757, 4.7 ppm) and an autolytic peptide of

trypsin ([M+H]$^+_{calc}$ m/z 3211.474, 2.2 ppm, residues 160-190) as well as a cross-linked product between CaM residues 127-148 and melittin residues 23-24 ([M+H]$^+_{calc}$ m/z 2774.276, 2.9 ppm), which was also detected with one methionine residue oxidized ([M+H]$^+_{calc}$ m/z 2790.271, 0 ppm). The monoisotopic masses of the cross-linked products (left to right) in this spectrum are: [M+H]$^+_{calc}$ 1731.876, 2202.165, 2483.303, 2701.365, 2774.268, 2790.271, 3144.486, 3160.481, 3212.757. For detailed information on the cross-linked products see **Table 2-1**. Peptides derived from melittin: residues 1-7 ([M+H]$^+_{calc}$ 675.429), 8-22 ([M+H]$^+_{calc}$ 1511.920), and 8-23 ([M+H]$^+_{calc}$ 1668.021). Peptides from calmodulin: residues 31-37 ([M+H]$^+_{calc}$ 805.424), 14-21 ([M+H]$^+_{calc}$ 956.472), 1-13 with acetylated N-terminus ([M+H]$^+_{calc}$ 1563.754), 91-106 ([M+H]$^+_{calc}$ 1754.871), 14-30 ([M+H]$^+_{calc}$ 1844.891), 107-126 with trimethyllysine K115 ([M+H]$^+_{calc}$ 2401.174 and 2417.169 (1x MSO) and 2433.163 (2x MSO)), 127-148 ([M+H]$^+_{calc}$ 2490.080 and 2506.075 (1x MSO) and 2522.070 (2x MSO)), 38-74 ([M+H]$^+_{calc}$ 4085.841 (1x MSO) and 4101.836 (2x MSO) and 4117.831 (3x MSO)). Peptides derived from trypsin autolysis: [M+H]$^+_{calc}$ 1153.574, 2163.056, 2273.159, 2289.154, and 3211.474.

Signals of CaM and melittin as well as signals of trypsin autolytic peptides are visible, with signals of the cross-linked products displaying only low intensities. The magnified inserts demonstrate the high resolution and mass accuracy provided by FTICR mass spectrometry. For example, the cross-linked product at m/z 3212.742 is clearly resolved from the signal at m/z 3211.467 belonging to an autolytic fragment (sequence 160-190, M180 oxidized) of trypsin. The mass accuracies for both signals were 4.7 ppm and 2.2 ppm, respectively; the average mass accuracy in the cross-linking experiment was 1.7 ppm using cross-linking reagents EDC / sNHS and trypsin as protease.

2.3.2.2 Cross-Linked Products Obtained with sDST

In addition to EDC / sNHS, CaM and melittin were cross-linked with the homobifunctional, amine-reactive cross-linker sDST (**Table 2-2**). After tryptic digestion of the CaM / Mel (1:1) complex three intermolecular cross-linked products were observed, one of them containing an additional modification by a partially hydrolyzed cross-linker, and three peptides were found to be modified by a hydrolyzed cross-linker alone (**Table 2-2**). Peptides modified by a hydrolyzed cross-linker do not provide any direct structural information on protein interfaces, but can yield important information on solvent accessibility of certain amino acids.

Table 2-2: Cross-Linked Products of CaM and Melittin with the Cross-Linking Reagent sDST. Tryptic cleavage. 50-fold excess of sDST over proteins. Reaction times (in minutes) are given in numbers (5 to 120), OH-sDST: partially hydrolyzed sDST.

CaM	Modified Residue	Melittin	Modified Residue	[M+H]$^+_{calc}$	[M+H]$^+_{exp}$	ppm	Sample	
1-21	K13	22-24	K23	3073.511	3073.529	6	50-5	**ORIENTATION A**
					3073.491	7	50-15	
					3073.503	3	50-30	
					3073.485	9	50-60	
					3073.532	7	50-120	
1-37 +OH-sDST	K13, K21, K30	22-24	K23	4896.337	4896.327	2	50-5	
					4896.385	10	50-15	
76-90	K77	8-22	K21	3636.858	3636.864	2	50-15	**ORIENTATION B**
					3636.878	6	50-60	
					3636.869	3	50-120	
-	-	1-7 +OH-sDST	G1	789.435	789.433	3	50-15	**HYDROLYZED LINKER**
14-30 +OH-sDST	K21	-	-	1976.897	1976.894	2	50-120	
75-86 +OH-sDST	K75, K77	-	-	1628.696	1628.701	3	50-15	

Two cross-linked products are consistent with orientation A of melittin in the complex (**Figure 1.9**). Cross-linking between K13 of CaM (amino acid sequence 1-21) and K23 of melittin within amino acid sequence 22-24 (signal at *m/z* 3073.511) was observed for all five incubation times.

Only one cross-linked product consistent with orientation B (**Figure 1.9**) was observed: the product between residues 76-90 of CaM and residues 8-22 of melittin was observed at *m/z* 3636.858 in three samples, pointing to K77 of CaM being cross-linked to K21.

In summary, for sDST only cross-linked products were detected, which connect melittin with the N-terminal domain and the central helix of CaM, but no cross-linking between melittin and the C-terminal domain of CaM was observed.

2.3.2.3   Cross-Linked Products Obtained with BS^3

Table 2-3 summarizes the cross-linked products obtained with BS^3 after tryptic digestion of the CaM / Mel (1:1) complex. As in the case of the experiments with sDST, two cross-linked products were consistent with orientation A, whereas one was consistent with orientation B (**Figure 1.9**).

The cross-linked product between residues 14-30 of CaM and residues 23-24 of melittin (signal at *m/z* 2285.166) is indicative of a cross-link between K21 of CaM and K23 of melittin and was observed once (**Table 2-3** sample 50-60, 50-fold excess of BS^3, 60 minutes reaction time). In addition, residues 76-90 of CaM and residues 1-7 of melittin were found to be cross-linked (signal at *m/z* 2666.334) via K77 of CaM and the N-terminus of melittin.

One cross-linked product pointing to orientation B was observed with a signal at *m/z* 2639.381, which was assigned to a cross-linked product between residues 14-30 of CaM and 1-7 of melittin. In this instance, K21 of CaM was cross-linked with the free N-terminus of melittin.

Table 2-3: Cross-Linked Products of CaM and Melittin with the Cross-Linker BS^3. Tryptic cleavage. 50-fold excess of BS^3 over proteins. Reaction times (in minutes) are given in numbers (5 to 120).

CaM	Modified Residue	Melittin	Modified Residue	[M+H]$^+_{calc}$	[M+H]$^+_{exp}$	ppm	Sample	
14-30	K21	23-24	K23	2285.166	2285.168	1	50-60	**ORIENTATION A**
76-90	K77	1-7	G1	2666.334	2666.334	0	50-15	
14-30	K21	1-7	G1	2639.381	2639.383	1	50-60	**ORIENTATION B**
					2639.363	7	50-120	
14-30 +OH-BS^3	K21	-	-	2000.970	2000.968	1	50-30	**HYDROLYZED LINKER**
					2000.968	1	50-60	
					2000.960	5	50-120	
75-90 +OH-BS^3	K75, K77	-	-	2785.360	2785.339	8	50-60	

As in the case of sDST, only BS^3-cross-linked products of melittin with the N-terminal domain of CaM and the central helix of CaM were observed, but none between melittin and the C-terminal domain of CaM. The reason why no cross-linked products between melittin and the C-terminal

lobe of CaM were observed with neither sDST nor BS[3] is probably due to the fact that there are only two lysine residues in the C-terminal domain of CaM. One is K115, which is trimethylated, the other one is K148, which represents the C-terminal amino acid of CaM.

2.4 Structural Model of the CaM / Melittin Complex

Figure 2.8 provides a schematic overview of the combined results from cross-linking experiments between melittin and CaM using the three cross-linking reagents EDC / sNHS, sDST, and BS[3] under the conditions described in *chapter 2.3*.

On top of the schematic melittin peptide all the amino acid residues of CaM, which are potentially involved in cross-linking pointing to orientation A, are displayed. Below the melittin peptide, the cross-linked amino acid residues of CaM are summarized that point to orientation B. Essentially, only three cross-linked sites in melittin were observed, namely the free N-terminus and lysines 21 and 23 in the C-terminal basic cluster. The cross-linked amino acid residues in CaM can be grouped into three categories: the N-terminal domain, the central helix, and the C-terminal domain of CaM. However, one specific amino acid of melittin can cross-link to both the N-terminal and the C-terminal domain of CaM in the same cross-linked complex species, i. e. with the same orientation of melittin within CaM. This is due to the fact that CaM wraps around the peptide, whereby the N- and C-terminal domains come closely together as described in *chapter 1.5.2*.
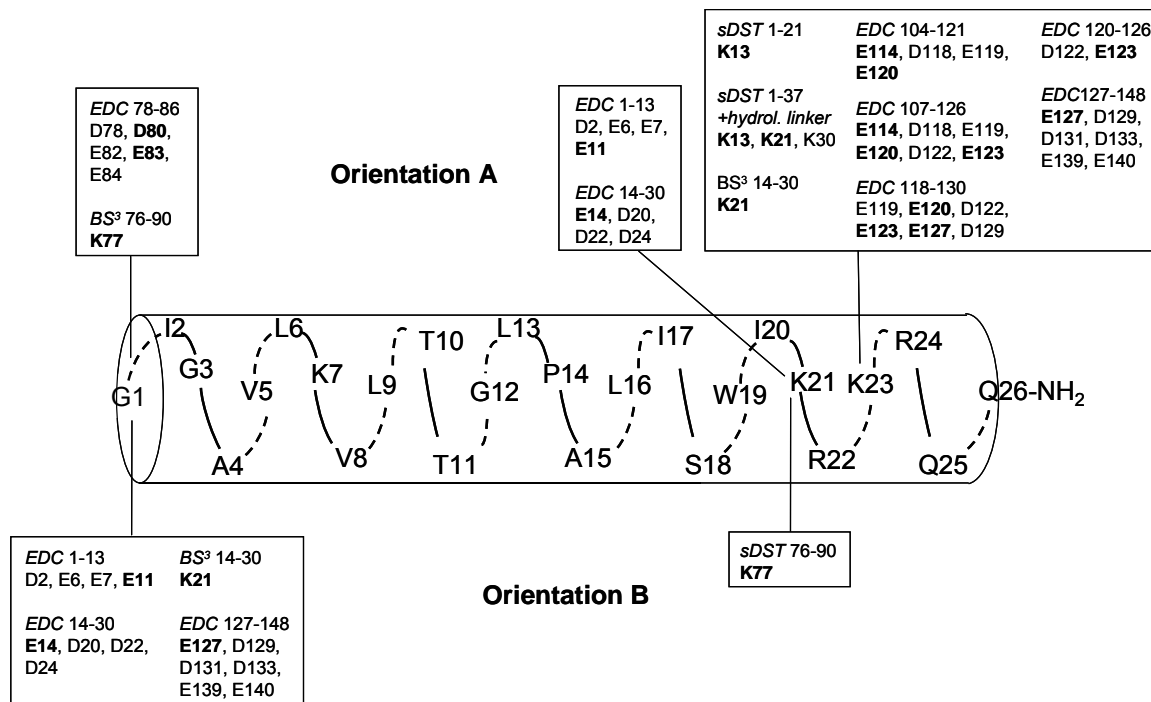
Figure 2.8: Schematic Model of Cross-Linked Residues Between Melittin and CaM. The melittin peptide is displayed in the center as α-helix. Cross-linked residues of CaM are located in the C- or N-terminal domain or in the central linker and can be assigned to either orientation A or orientation B. Amino acids, which are most likely to be involved in cross-linking, are printed in bold.

Low-resolution structural models were calculated from ambiguous distance restraints derived from the cross-linking data using conjoined rigid body / torsion angle simulated annealing. The backbone and non-interfacial side chains of the N- and C-terminal domains of CaM were treated as rigid bodies connected by a linker with full torsional degrees of freedom. Likewise, the two helical halves of melittin were treated as rigid bodies connected by a linker (residues 10-13) with torsional degrees of freedom. All interfacial side chains were given torsional degrees of freedom. The distance restraints are represented as an empirical $<r^{-6}>^{-1/6}$ average and allow for both ambiguity in terms of cross-linking assignments and melittin orientation. The two orientations of bound melittin were calculated simultaneously. A ribbon diagram of a representative pair of structures with opposite orientations of melittin in the CaM / Mel complex is displayed in **Figure 2.9**. The principal limitation in generating higher resolution structures from the cross-linking data is due to the presence of only three hinge residues for melittin (K21, K23, and the N-terminus).
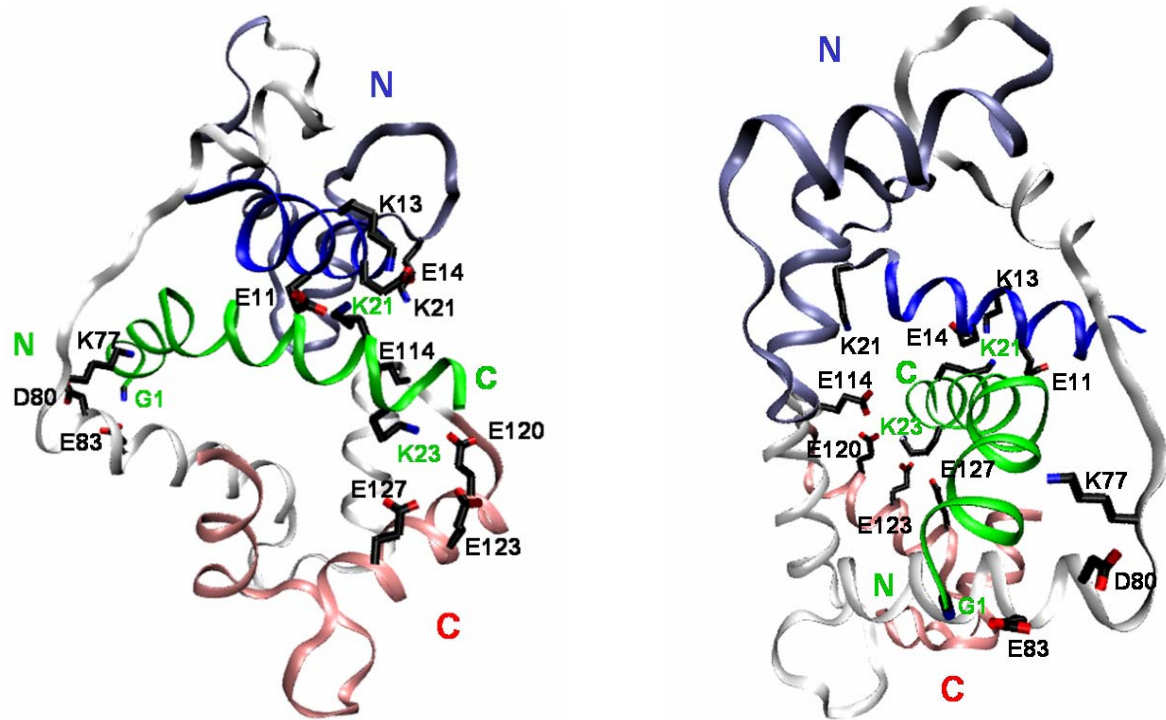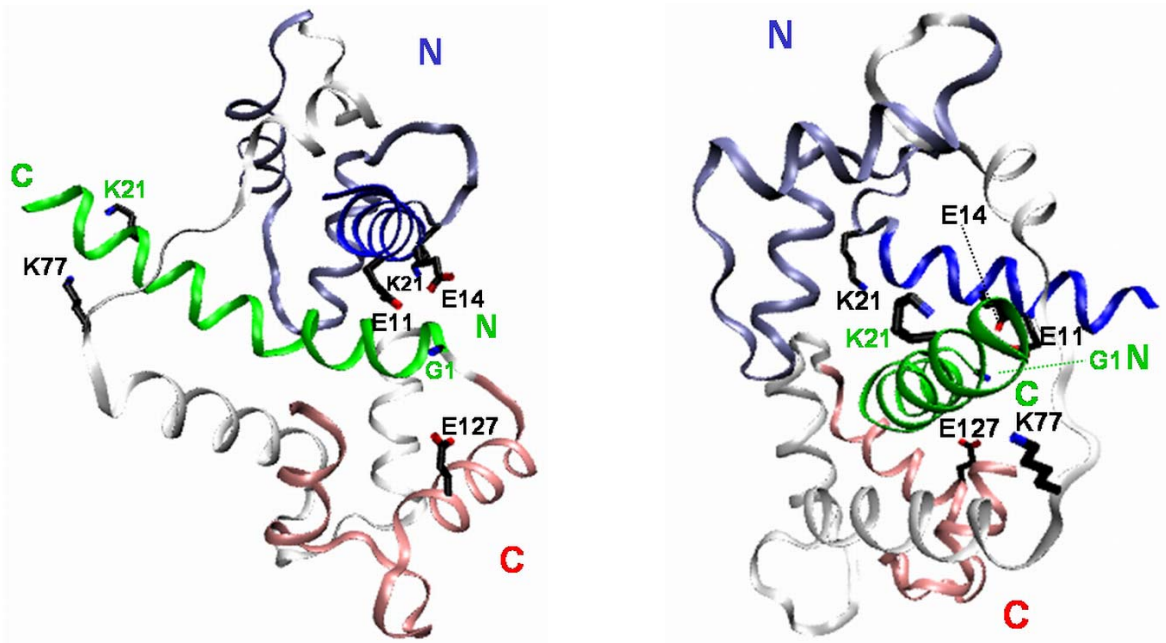
Figure 2.9: Low-Resolution Models of the CaM / Mel Complex. Parallel (A) and anti-parallel (B) modes of binding of melittin in the CaM-melittin complex calculated from ambiguous distance restraints derived from the cross-linking data using conjoined rigid body/torsion angle simulated annealing. The amino acid sidechains of CaM and melittin that are involved in cross-linking are shown for both orientations A and B. Each orientation is represented by two views: one with the helix of melittin parallel (*left*) and one with the helix of melittin perpendicular (*right*) to the plane of the page.

2.5 Discussion

2.5.1 Chemical Cross-Linking and Mass Spectrometry for Studying Mechanisms of Protein-Protein Interactions

As exemplified for the CaM / Mel complex, chemical cross-linking techniques combined with mass spectrometry and appropriate computational methods provide the basis for an efficient low-resolution three-dimensional structural characterization of protein complexes.

The implication of mass spectrometric analysis for this kind of protein-protein interaction studies is apparent because of its high sensitivity enabling rapid analysis of the complex mixtures obtained from enzymatic digests of cross-linking reaction mixtures. If one can obtain a sufficient number of cross-links, these can be used as distance restraints – albeit ambiguous – to determine a low-resolution structure of a protein / peptide or protein / protein complex, provided that high resolution structures of the involved interaction partners are available. In this regard, it was shown that this approach represents a low-resolution alternative to NMR spectroscopy and X-ray crystallography-based methods for low-resolution structural studies of macromolecular complexes.

2.5.2 Structures of CaM / Peptide Complexes

The structures of the CaM / peptide complexes solved to date [Ikura *et al*., 1992, Osawa *et al*., 1999, Meador *et al*., 1992 & 1993]   had identified several important features for target peptide binding to CaM as described above. **In Figure 1.9** several target peptide sequences, as well as their respective orientations, within CaM / peptide complexes are displayed. In most previously solved CaM / target peptide complexes, such as CaM complexed with skMLCK peptide [Ikura *et al*., 1992], smMLCK [Meador *et al*., 1992], and CaMKII [Meador *et al*., 1993], the peptide α-helix is positioned in a way that the N-terminal portion of the peptide interacts with the C-terminal domain of CaM, whereas the C-terminal portion binds to the N-terminal domain of CaM (orientation B, **Figure 1.9**). The peptide's orientation is reversed in the case for CaM-CaMKKα complexes [Osawa *et al*., 1999, Kurokawa *et al*., 2001] (orientation A **Figure 1.9**).

Recently, two novel structures of CaM complexed with large target fragments, namely the calcium activated K^+ channel [Schumacher *et al*., 2001] and the anthrax exotoxin [Drum *et al*., 2002], have been solved and revealed unexpected binding modes. This included dimer formation with the fragments of the K^+ channel and insertion of CaM between two domains of the anthrax toxin. These reports suggests that structural data on CaM complexed with large target proteins yield additional information on the mechanism of CaM target recognition and demonstrate that CaM possesses a variety of interaction modes with its targets.

2.5.3   Structure of the CaM / Melittin Complex

To date, there has only been one structural study of the CaM / Mel complex, based on limited proteolysis and chemical cross-linking experiments with EDC, [Scaloni *et al.*, 1998]. In this previous study, an exclusive orientation A of melittin in the complex was suggested based on the identification of only two cross-linked products connecting the N-terminus and K23 of melittin with residues 77-84 and E11, respectively, of CaM [Scaloni *et al.*, 1998]. The structural investigation presented here confirms and extends these findings by detecting cross-linking products between the N-terminus of melittin and residues 78-86 as well as between K21 of melittin and residues 1-13 of CaM (**Table 2-1**). Another report on the CaM / Mel complex had suggested that 80% of the melittin population binds to CaM in the orientation A, however, no data were shown supporting this statement [Brokx *et al.*, 2001].

To generate low resolution models of the CaM / Mel complex on the basis of the experimental cross-linking data conjoined rigid body / torsion angle simulated annealing [Clore, 2000, Schwieters & Clore, 2001, Clore & Bewley, 2002] was employed, a technique that has proven very powerful for solving structures of macromolecular complexes from NMR data [Wang *et al.*, 2000, Cornilescu *et al.*, 2002, Cai *et al.*, 2003]. The limitation in generating structures from the cross-linking data is due to the presence of only three hinge residues, K21, K23, and the N-terminus, for melittin. K7, which could potentially form cross-linked products with CaM, was never found to be cross-linked in our experiments. The most likely reason is that cross-linked products involving K7 are not amenable to detection under the mass spectrometric conditions employed, as the proteases used are unable to cleave after K7 is it is modified.

From a qualitative examination of the cross-linking data (**Tables 2-1, 2-2, and 2-3**), it became immediately apparent that the data were not compatible with a unique orientation of melittin in the CaM / Mel complex. Structures were generated from the cross-linking data ascribing melittin to two opposite orientations within the CaM / peptide complex (**Figure 2.9**), one parallel (orientation A) and the other anti-parallel (orientation B). The majority of cross-links originate from orientation A (**Figure 2.8**). This could have been anticipated from the basic amino acid cluster (K21-K24) located close to the C-terminus of melittin. However, there are a number of cross-linking derived distance restraints that are not satisfied in the structures within the distance limits given by spacer lengths of the respective cross-linkers. One possible explanation is that even within the context of two opposing orientations of melittin, multiple modes of binding exist and that cross-linking may pick up minor conformational states due to its fast reaction kinetics. However, it is problematic to quantify the distribution of orientations of a protein/peptide in the complex from the number of cross-links obtained. Certain cross-linked products might be created in the course of the reaction, but are not amenable to mass spectrometric detection. During the cross-linking reaction, a variety of different complexes are created but only those that

can be picked up during mass spectrometric detection are considered for building a structural model.

The finding that melittin can bind to CaM in two opposite directions suggests that the presence of the basic amino acid cluster in the target peptide is not the only crucial determinant of peptide orientation within the CaM / Mel complex. It was demonstrated that CaM is able to recognize its target peptides in multiple modes, which is also evidenced by the recently published structures of complexes between CaM and the larger target proteins mentioned above [Schumacher 2001, Drum 2002]. It seems that the two hydrophobic anchor residues in melittin, L6 and W19, can be accommodated in each of the hydrophobic patches, and that the bulky amino acid W19 does not necessarily have to be accommodated in the C-terminal domain of CaM. Apparently, the basic amino acid cluster in melittin, comprising amino acids K21 to R24, can interact with negatively charged amino acids (glutamic or aspartic acids) in either of CaM's CO-2 or CO-1. From the number of cross-linked products obtained for each orientation, it is likely that melittin binds predominantly in the parallel orientation A.

3 Purification and Characterization of the Annexin A2 / P11 Complex

For structurally characterizing the heterotetrameric complex between annexin A2 and p11 (A2t) by chemical cross-linking and mass spectrometry, the complex was purified from mucosa of pig (*Sus scrofa*) small intestines. The basic principle underlying the purification of A2t is the Ca^{2+}-induced reversible association of the complex to phospholipid membranes. Taking advantage of this mechanism, the purification protocol involves first the attachment of A2t to the membrane by adding Ca^{2+} and removal of undesirable components, before A2t is released from the membrane by adding the chelator ethylene glycol tetraacetic acid (EGTA). The purified complex was identified and characterized with respect to amino acid sequences, molecular weight, and post-translational modifications of the subunits.

3.1 Purification of A2t from Pig Small Intestine

A2t was purified from porcine small intestinal mucosa following a slightly modified version of the protocol developed by Gerke and Weber [1984] as described in chapter 6.2.

3.1.1 Preparation of Small Intestines

Six fresh pig small intestines were obtained from the slaughterhouse Altenburg and processed immediately after exvisceration. The gut contents were removed by washing the intestinal lumen with ice-cold buffer. Afterwards, the small intestines were trenched into pieces of about 50 cm length and were slit lengthwise. The interior thin mucosal layer was scraped off and was immediately frozen in liquid nitrogen for storage.

The small intestine is divided into three parts: duodenum, jejunum, and ileum. **Figures 3.1 A** and **B** depict the luminal surface and a section of the intestinal wall of the ileum, respectively.

A

B
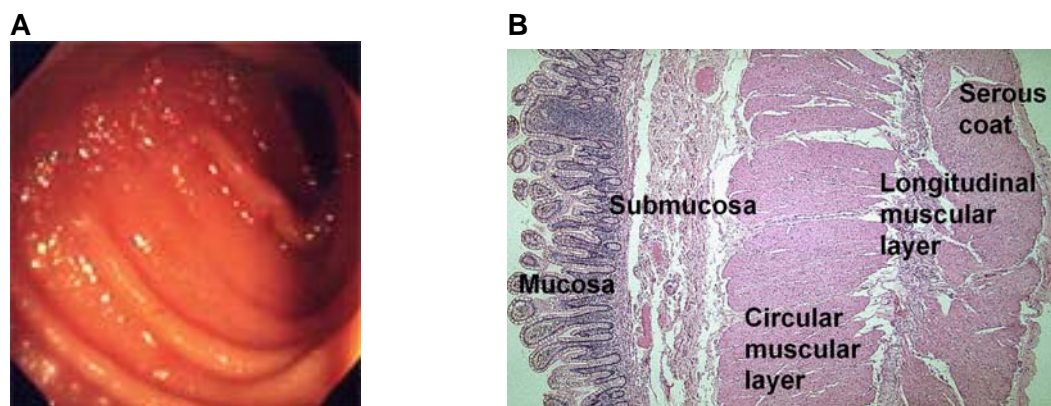
Figure 3.1: Small Intestine. *A*, Luminal surface of the ileum. The mucosa appears as a shiny layer of mucus. *B*, Section through the ileal wall. The different tissue types are indicated.

For successful protein purification it was critical to scrape off only the thin mucosa layer, but not the subjacent tissue.

3.1.2 Sample Preparation for Chromatographic Separation

All steps of the purification procedure were carried out at 4°C. Protease inhibitors and dithiothreitol (DTT) were applied during the course of the purification to avoid sample degradation and aggregation of p11 and ANXA2 by oxidation of free cysteine thiols to disulfides.

For preparation of the sample for the subsequent chromatographic separation, the frozen material was thawed, buffer was added, and the mucosa was mechanically disrupted. In addition to protease inhibitors the buffer contained Ca^{2+}, which causes the ANXA2 / p11 complex to bind to the membrane. Six washing and centrifugation steps were conducted, during which the supernatant was discarded and the pellet containing A2t was retained and resuspended in buffer. After each step, aliquots were taken for monitoring the progress of purification. **Figure 3.2** illustrates the effect of the repeated washing steps. Lanes 1 through 6 display the supernatant aliquots. Note that the samples contained in lanes 4, 5, and 6 are ten-fold concentrated, demonstrating efficient removal of undesirable soluble proteins and other tissue components. Lane 8 represents the resuspended pellet after the washing procedure, containing membrane-bound A2t and other membrane-associated proteins as well as insoluble tissue components.
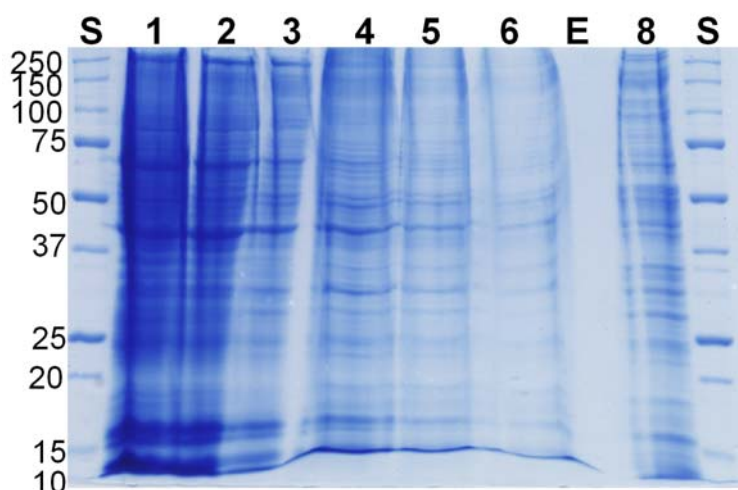
Figure 3.2: SDS-PAGE of the Washing Procedure. S: molecular weight standard (kDa), lanes 1, 2, and 3: supernatants of washing steps 1, 2, and 3, respectively; washing buffer contained Triton X-100. Lanes 4, 5, and 6: supernatants of washing steps 4, 5, and 6, respectively, 10x conc.; washing buffer without Triton X-100. Lane 8: resuspended pellet, E: empty lane; 5% stacking / 12% resolving gel.

Next, the pellet was resuspended in EGTA-containing buffer, and the cells were thoroughly disrupted by shear force in a Potter homogenizer. Complexation of Ca^{2+} by EGTA causes A2t to be released from the membrane. After ultracentrifugation the A2t containing supernatant was retained for subsequent purification by ion exchange chromatography.

3.1.3 Protein Purification by Fast Protein Liquid Chromatography

Prior to protein purification by diethylaminoethyl cellulose (DEAE) ion exchange chromatography, the supernatant obtained from ultracentrifugation was dialyzed over-night against buffer containing 20 mM imidazole, 0.5 mM EGTA, 1 mM DTT, pH 7.5, 10 mM NaCl. The dialyzed protein solution was applied onto the equilibrated DEAE column and fast protein liquid chromatography (FPLC) was conducted. The flow rate was manually adjusted and varied from 2-4 ml/min. Earlier purification test runs indicated that applying a two-step NaCl gradient (10 mM and 1 M) is sufficient for an efficient separation of A2t from accompanying proteins (**Figure 3.3**).
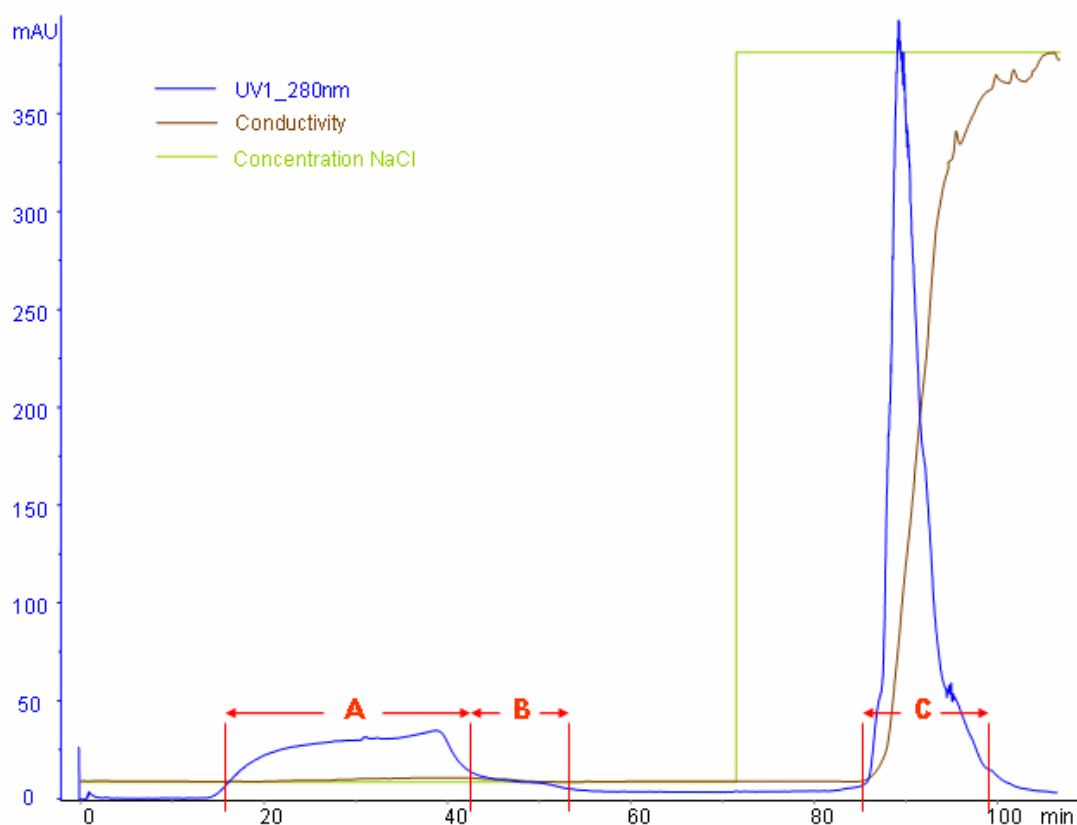
Figure 3.3: Chromatogram of A2t Purification by Anionic Exchange Chromatography.
Fractions A, B, and C were collected.

The fractions were collected manually based on the UV absorption at 280 nm. Due to negligible interactions of A2t with the DEAE material at pH 7.5, A2t was obtained in the flow-through fractions A and B (**Figure 3.3**) at the starting NaCl concentration of 10 mM. Fraction C was

obtained at 1M NaCl and comprises those proteins that bound to the column. Aliquots of each of the three fractions were subjected to SDS-PAGE analysis (**Figure 3.4**), with lanes 1, 2, and 3 corresponding to chromatographic fractions A, B, and C, respectively. A2t that was obtained in fraction A (lane 1), which contains minor contaminants between 15 and 20 kDa. Fraction B (**Figure 3.3**, B, **figure 3.4**, lane 2) , which was collected after an abrupt drop in absorption, resembles fraction A. Fraction C contains the undesired proteins (lane 3) as well as ANXA2 and p11 at low concentrations (lane 3). Thus, the gel clearly demonstrates the conceptually simple, yet highly efficient purification of A2t by a single chromatographic separation step.
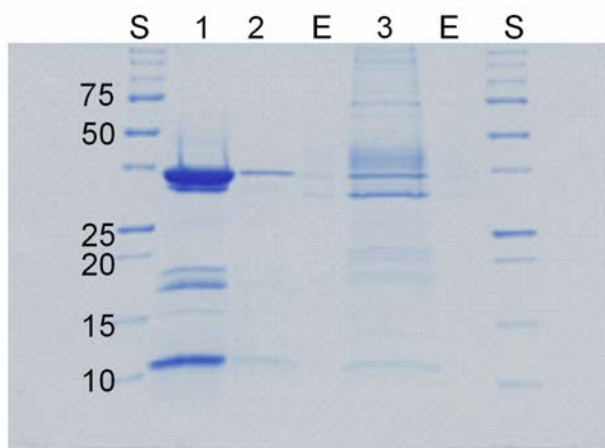
Figure 3.4: SDS-PAGE of Fractions from Anionic Exchange Chromatography of A2t. Lanes 1, 2, and 3 represent fractions A, B, and C, respectively. S: molecular weight standard (kDa), E: empty lane.

A2t was about 90% pure (SDS-PAGE) after anionic exchange chromatography and was used without further purification by strong cationic exchange chromatography (SCX) as had been suggested in the original protocol [Gerke & Weber 1984, Menke, 2005], because the remaining impurities could not be removed by SCX (data not shown).

Fraction A containing purified A2t was concentrated using Centriprep centrifugal filter devices. The buffer was at the same time exchanged to an appropriate buffer for subsequent chemical cross-linking experiments. In total, 12 mg protein were obtained from ca. 900 g mucosal scrapings. A2t was lyophilized for storage.

3.1.4 Mass Spectrometric Identification of ANXA2, P11, and Co-Purified Proteins

ANXA2, p11, and accompanying proteins of A2t were identified by peptide mass fingerprint analysis (PMF). Furthermore, some proteins of the solution that was loaded onto the DEAE column and a number of proteins that eluted with 1M Nail (fraction C) were analyzed.

Following gel electrophoretic separation of aliquots of the sample loaded onto the column and of the different stages of the chromatographic separation, the highlighted gel bands (**Figure**

3.5) were excised and subjected to enzymatic proteolysis. The peptide mixtures from tryptic *in-gel* digestion were analyzed by nano-HPLC / nano-ESI-FTICRMS as well as by MALDI-TOFMS.
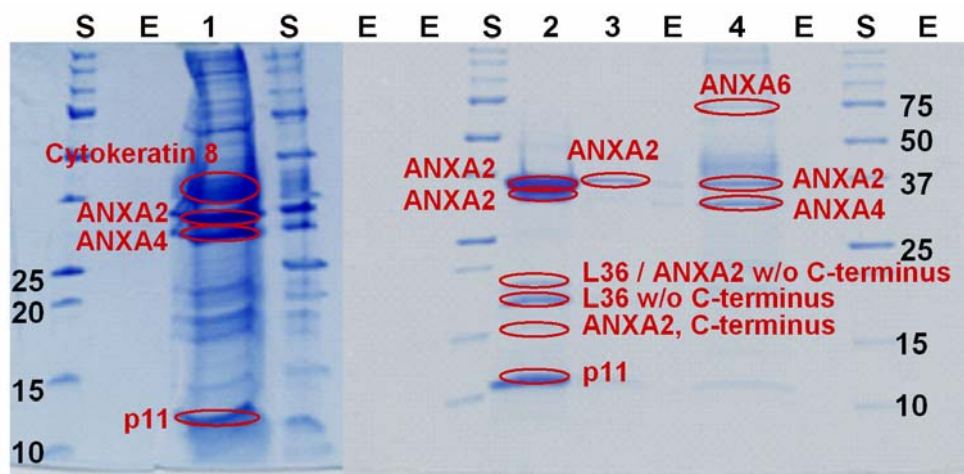
Figure 3.5: **SDS-Polyacrylamide Gel for Peptide Mass Fingerprint Analysis by MALDI-TOFMS**. S: molecular weight standard (kDa), lane 1: sample loaded on DEAE column, dialyzed, lane 2: DEAE flow through fraction A, lane 3: DEAE flow through fraction B, lane 4: Fraction C eluted with 1M NaCl, E: empty lane. 15% gel.

As expected, ANXA2 and p11 were identified in the DEAE flow-through (**Figure 3.3** fraction A, **Figure 3.5** lane 2, ~37 and 12 kDa), evidencing purification of A2t. However, minor impurities (bands migrating between ~ 15 and 25 kDa, **Figure 3.5** lane 2) are still present. These impurities turned out to be degradation products of ANXA2 with missing C-terminus as well as lactose-binding protein L36 (galectin). Tomonaga *et al.* (2004) reported on C-terminally truncated ANXA2 species due to *in-situ* proteolysis in healthy colon, which was obtained as a gel band migrating at around 25 kDa and confirmed by mass spectrometry. The protein band migrating just below the ANXA2 band (**Figure 3.5**, lane 2) was also identified to be ANXA2, but turned out to be N-terminally truncated as will be detailed in *chapter 3.2.1*. Apart from ANXA2 (38.5 kDa), other annexins, such as ANXA4 (36 kDa) and ANXA6 (70 kDa) (**Figure 3.5** lane 4), were identified, proving the selective isolation of calcium-dependent membrane binding proteins in the preceding purification steps. ANXA4 and ANXA6 possess pI values of 5.7 and 5.4, respectively, and therefore exhibit an overall negative charge at pH 7.5, enabling interaction with the DEAE column. Cytokeratin 8 (42.5 kDa, **Figure 3.5**, lane 1) is a ubiquitous component of membranes.

Comparison of the sample loaded onto the DEAE column (**Figure 3.5** lane 1) with the flow-through fraction A (**Figure 3.5** lane 2) also demonstrates the efficient one-step chromatographic separation of A2t. Apparently, undesired proteins bound to the column as did a small portion of A2t, which was eluted with 1M NaCl (lane 4).

Figure 3.6 shows the MALDI-TOF mass spectrum of a tryptic digest of ANXA2 (**Figure 3.5**, lane 2, ~37 kDa). For PMF the NCBI database was searched for peptides matching the obtained *m/z* values using the Mascot search engine. For MALDI-TOFMS data, a maximum

mass deviation of 20-40 ppm was allowed as well as one missed cleavage. Cysteines were defined to be carbamidomethylated and methionine oxidation (MSO) was specified as variable modification.
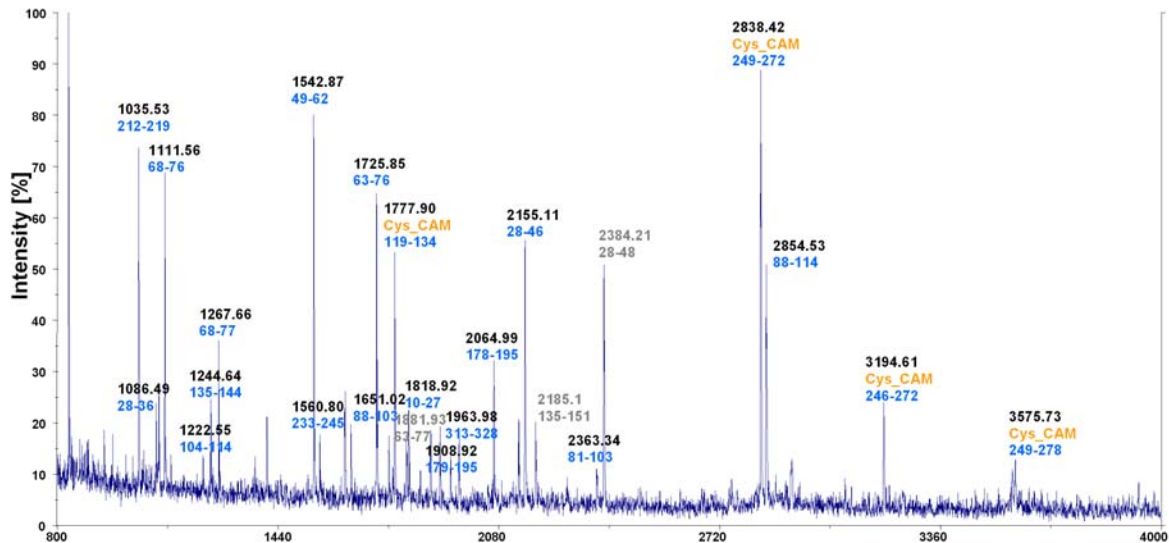
Figure 3.6: MALDI-TOF Mass Spectrum for Peptide Mass Fingerprint Analysis of Annexin A2. The signals were assigned to peptides of annexin A2 by *Mascot*. Cys_CAM: carbamidomethylated cysteine, blue: found when searching without MSO; gray: oxidized methionine (MSO).

21 peptides with an average mass deviation of 16 ppm were matched to ANXA2 from pig with a MOWSE score of 315 (**Figure 3.7**).
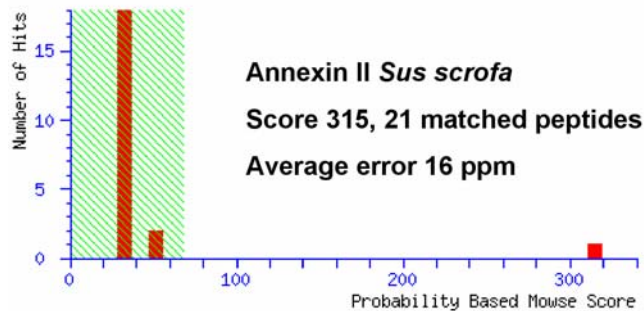
Figure 3.7: Identification of Annexin A2.

The sequence coverage of ANXA2 from the single MALDI-TOF mass spectrum presented in **Figure 3.6** was 63 %.

3.2 Characterization of A2t

Precise information about the proteins that comprise the complex to be studied is indispensable for a successful subsequent structural characterization. A2t was scrutinized with respect to the

molecular weights, amino acid sequences and potentially present post-translational modifications of ANXA2 and p11.

3.2.1   Gel Electrophoresis

From SDS gel electrophoresis bands migrating at ~37 kDa (ANXA2) and ~11 kDa (p11) were observed (**Figure 3.4**). Furthermore, a protein band migrating just below the ANXA2 band was observed, which might represent an N-terminally truncated form of ANXA2.

From peptide mass fingerprint analysis of tryptic and GluC digests it was concluded that the N-terminal amino acids 1-27 are apparently missing in the band migrating at around 33 kDa. **Figure 3.8** depicts the sequence coverage from tryptic digestion of the two ANXA2 bands.

```
Annexin A2 – Tryptic Digestion
```

```
AcSTVHEILCKL  SLEGDHSTPA  SAYGSVKAYT  NFDAERDALN  IETAIKTKGV  DEVTIVNILT 60
   NRSNEQRQDI  AFAYQRRTKK  ELASALKSAL  SGHLETVILG  LLKTPAQYDA  SELKASMKGL 120
   GTDEDSLIEI  ICSRTNQELQ  EINRVYKEMY  KTDLEKDIIS  DTSGDFRKLM  VALAKGRRAE 180
   DGSVIDYELI  DQDARDLYDA  GVKRKGTDVP  KWISIMTERS  VCHLQKVFER  YKSYSPYDML 240
   ESIKKEVKGD  LENAFLNLVQ  CIQNKPLYFA  DRLYDSMKGK  GTRDKVLIRI  MVSRSEVDML 300
   KIRSEFKRKY  GKSLYNYIQQ  DTKGDYQKAL  LYLCGGDD 338
```

```
AcSTVHEILCKL  SLEGDHSTPA  SAYGSVKAYT  NFDAERDALN  IETAIKTKGV  DEVTIVNILT 60
   NRSNEQRQDI  AFAYQRRTKK  ELASALKSAL  SGHLETVILG  LLKTPAQYDA  SELKASMKGL 120
   GTDEDSLIEI  ICSRTNQELQ  EINRVYKEMY  KTDLEKDIIS  DTSGDFRKLM  VALAKGRRAE 180
   DGSVIDYELI  DQDARDLYDA  GVKRKGTDVP  KWISIMTERS  VCHLQKVFER  YKSYSPYDML 240
   ESIKKEVKGD  LENAFLNLVQ  CIQNKPLYFA  DRLYDSMKGK  GTRDKVLIRI  MVSRSEVDML 300
   KIRSEFKRKY  GKSLYNYIQQ  DTKGDYQKAL  LYLCGGDD 338
```

Figure 3.8: Putative N-terminal Truncation of Annexin A2. Upper sequence: Upper ANXA2 band at ~ 37 kDa, blue amino acids were covered (69.5%) by tryptic digestion. Lower sequence: Lower ANXA2 band at ~ 33 kDa, red amino acids were covered (64.8%) by tryptic digestion. Ac: N-terminal acetylation.

In the upper ANXA2 band at ~ 37 kDa, the N-terminus was fully covered, whereas amino acids 1-27 were missing in the faster migrating ANXA2 band at ~ 33 kDa. For both gel bands the C-terminal amino acids 324-338 were not observed. Therefore, a C-terminal truncation rather than an N-terminal truncation cannot be repudiated. For the GluC digest, the same result was observed (data not shown) but, the sequence coverage was lower compared to the tryptic digest.

The proteolytic truncation of ANXA2 has been described by other groups [Glenney *et al.*, 1985, Johnsson *et al.*, 1986, Waisman 1995] and probably represents an artefact created during protein purification. Danielsen *et al.* (2003) observed a 33 kDa-annexin A2 species, which is probably generated from full-length ANXA2 by *in-situ* proteolytic cleavage caused by

pancreatic proteases as soon as ANXA2 is secreted to the luminal side of the mucosa's brush border membrane.

Native gel electrophoresis at 4°C (**Figures 3.9 A**, lane 3 and **3.9 B**, lane 1) without SDS and reducing agents confirmed purification of the intact non-covalent complex between ANXA2 and p11. The purified complex appeared on the gel as a broad band migrating between ca. 140 and 300 kDa, which indicates oligomer formation. The existence of an octameric state of the ANXA2 / p11 complex has been proposed by several groups [Waisman, 1995, Lambert *et al.*, 1997]. However, polyacrylamide gel electrophoresis is inadequate for an exact molecular weight determination. Nevertheless, the occurence of the high-molecular weight protein band hints towards oligomer formation.
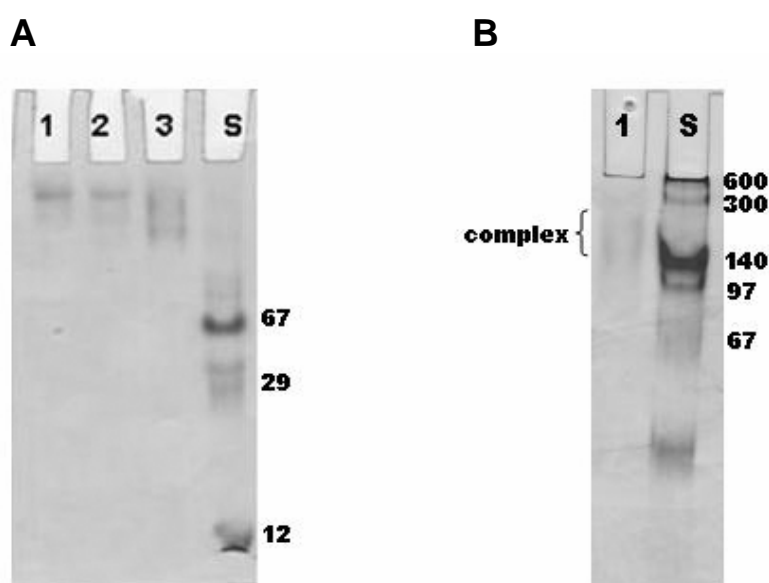
Figure 3.9: Native Gel Electrophoresis of A2t. *A*, Native PAGE of cross-linked and non-cross-linked A2t. 8% continuous polyacrylamide gel; lanes 1 and 2: A2t cross-linked with sDST will be described in *chapter 4.1.1*, lane 3: non-cross-linked A2t, S: native molecular weight standard (mixture of cytochrome c (12 kDa), carbonic anhydrase (29 kDa), and BSA (67 kDa)). *B*, Native PAGE of non-cross-linked A2t. 8% continuous polyacrylamide gel; lane 1: non-cross-linked A2t, S: native molecular weight standard (mixture of BSA (67 kDa), phosphorylase b (97 kDa), lactate dehydrogenase (140 kDa), and IgA (300 kDa)).

3.2.2 Amino Acid Sequences of ANXA2 and p11

ANXA2 and p11 were scrutinized with respect to their amino acid sequences. The amino acid sequence of p11 is deposited in the Swiss-Prot database (www.expasy.org) (entry P04163), but, at present, there is no complete amino acid sequence for porcine ANXA2 available in the Swiss-Prot database. Only an N-terminal fragment of porcine ANXA2 (entry P19620) comprising amino acids 1-91 is currently deposited in the Swiss-Prot database. Luckily, the

Swiss-Prot database contains ANXA2 entries for a number of vertebrate species. Vertebrate annexins exhibit great sequence homologies and were therefore used as templates for scrutinizing the amino acid sequence of ANXA2 from *Sus scrofa*. Furthermore, complete porcine ANXA2 sequences suggested by Manuela Menke [Menke, 2005] and Volker Gerke (*unpublished*) as well as the N-terminal fragment (aa 1-91) of porcine ANXA2 were considered. **Figure 3.10** lists all the ANXA2 amino acid sequences that were compared.

```
STVHEILCKL SLEGDHSTPA SAYGSVKAYT NFDAERDALN IETAIKTKGV DEVTIVNILT⁶⁰
.......... .........A .......... .......... .......... ..........   Pig aa 1-91
.......... .........P .......... .......... .......... ..........   Bovine
.......... .........P .......... .......... .......... ..........   Human
.......... .........P .......... .......... .......... ..........   Canine
.......... .........A .......... .......... .......... ..........   VG, Pig
.......... .........A .......... .......... .......... ..........   MM, Pig


NRSNEQRQDI AFAYQRRTKK ELASALKSAL SGHLETVILG LLKTPAQYDA SELKASMKGL¹²⁰
....E.X..- ---------- ---------- ---------- ---------- ----------
....E.R... .......... .......... .......... ........D. ..........
....A.R... .......... .......... .......... ........D. ..........
....E.R... .......... .......... .......... ........D. ..........
....E.R... .......... .......... .......... ........N. ..........
....E.R... .......... .......... .......... ........N. ..........


GTDEDSLIEI ICSRTNQELQ EINRVYKEMY KTDLEKDIIS DTSGDFRKLM VALAKGRRAE¹⁸⁰
---------- ---------- ---------- ---------- ---------- ----------
.......... .......... .......... ........V. ......R... ..........
.......... .......... .......... ........I. ......R... ..........
.......... .......... .......... ........I. ......R... ..........
.......... .......... .......... ........I. ......A... ..........
.......... .......... .......... ........I. ......A... ..........


DGSVIDYELI DQDARDLYDA GVKRKGTDVP KWISIMTERS VCHLQKVFER YKSYSPYDML²⁴⁰
---------- ---------- ---------- ---------- ---------- ----------
.......... .......... .......... .......... .C......E. ..........
.......... .......... .......... .......... .P......D. ..........
.......... .......... .......... .......... .C......E. ..........
.......... .......... .......... .......... .C......E. ..........
.......... .......... .......... .......... .C......V. ..........


ESIKKEVKGD LENAFLNLVQ CIQNKPLYFA DRLYDSMKGK GTRDKVLIRI MVSRSEVDML³⁰⁰
---------- ---------- ---------- ---------- ---------- ----------
...K...... .......... .......... .......... .......... ..........
...R...... .......... .......... .......... .......... ..........
...K...... .......... .......... .......... .......... ..........
...R...... .......... .......... .......... .......... ..........
...R...... .......... .......... .......... .......... ..........


KIRSEFKRKY GKSLYNYIQQ DTKGDYQKAL LYLCGGDD³³⁸
---------- ---------- ---------- --------
.......K.. .....Y.... .......... ........
.......R.. .....Y.... .......... ........
.......R.. .....Y.... .......... ........
.......R.. .....N.... .......... ........
.......R.. .....N.... .......... ........
```

Figure 3.10: Comparison of ANXA2 Amino Acid Sequences from Different Vertebrate Species. The reconstructed porcine ANXA2 amino acid sequence is given on top. Identical amino acids are represented as dots. A dash represents a non-existent amino acid. VG: V. Gerke, MM: M. Menke.

The first line shows the porcine ANXA2 sequence as reconstructed from other vertebrate species. Presented below are the N-terminal porcine fragment (aa 1-91), the sequences

available for *Bos taurus* (entry P04272), *Homo sapiens* (entry P07355), and *Canis familiaris* (entry Q6TEQ7), and the sequences suggested by Volker Gerke (*unpublished*) and Manuela Menke [Menke, 2005]. All the amino acid sequences listed differ from each other by at least two amino acids.

The *m/z* values obtained from nano-HPLC / nano-ESI-FTICRMS analysis of tryptic digests were compared to all ANXA2 sequences presented in **Figure 3.10** using the ExPASy *FindPept* tool. Essentially, the amino acid sequences of bovine and human ANXA2 as well as the N-terminal porcine fragment were used for generating the final porcine ANXA2 sequence as they proved to cover most of the observed amino acid exchanges (**Figure 3.11**). However, the sequences suggested by V. Gerke and M. Menke led to the assumption of amino acid 316 being an asparagine instead of a tyrosine as observed for other vertebrate species. On the other hand, the amino acid sequences suggested by V. Gerke and M. Menke turned out to not completely match the reconstructed porcine ANXA2 sequence. In fact, the sequences differed by three and four amino acids, respectively. In case of the sequence proposed by M. Menke, only 27.2% of the suggested sequence was covered by MALDI-TOFMS. With the exception of N316, none of the uncertain amino acids were validated, as the corresponding peptide masses were not detected [Menke, 2005].

From the scrutinized vertebrate ANXA2 sequences, porcine ANXA2 shares the greatest sequence homology with ANXA2 of *Canis familiaris*, with only two amino acid exchanges. Porcine ANXA2 differs from human ANXA2 by six amino acids and from bovine ANXA2 by four amino acids (**Figure 3.11**).

Porcine	A20	E65	D109	I159	R167	C222	E229	K244	R308	N316
Pig (aa 1-91)	A	E	-	-	-	-	-	-	-	-
Bovine	P	E	D	V	R	C	E	K	K	Y
Human	P	A	D	I	R	P	D	R	R	Y
Canine	P	E	D	I	R	C	E	K	R	Y
VG, Pig	A	E	N	I	A	C	E	R	R	N
MM, Pig	A	E	N	I	A	C	V	R	R	N

Figure 3.11: Overview of Amino Acid Exchanges in Vertebrate ANXA2. Underlined amino acids differ from the sequence of purified ANXA2.

Having obtained the porcine ANXA2 amino acid sequence, the goal was to derive complete sequence coverages for ANXA2 and p11. For peptide mass fingerprint (PMF) analysis, trypsin, LysC and endoproteinase AspN were used for *in-gel* digestion, whereas chymotrypsin was additionally employed to attain improved sequence coverage for p11. The obtained peptide mixtures were analyzed by nano-HPLC / nano-ESI-FTICRMS as well as by MALDI-TOFMS. GPMAW and the ExPASy *FindPept* tool were employed for matching *m/z* values to amino acid

sequences. For ANXA2, trypsin alone yielded 72.5% sequence coverage, whereas AspN alone yielded 95.9%. The combined results from trypsin and AspN digests accounted for 98.2% of the ANXA2 sequence, with amino acids 271-275 (DRLY) and the C-terminal amino acids 337-338 (DD) still lacking. MS analyses of LysC digests of ANXA2 yielded a sequence coverage of 89.1%, covering the six missing amino acids. Thus, all the digests taken together provided a 100% sequence coverage of ANXA2 (**Figure 3.12**).

MS analyses of enzymatic digests of p11 yielded a sequence coverage of 95.8%. (**Figure 3.12**). Without chymotrypsin merely 71.6% of the sequence was covered, lacking the hydrophobic stretch of amino acids 57-83 that was not amenable for digestion by neither trypsin nor AspN. Chymotryptic digestion revealed most of the missing amino acids with the exception of amino acids 73-76 (**Figure 3.12** underlined).

Annexin A2

```
AcSTVHEILCKL  SLEGDHSTPA  SAYGSVKAYT  NFDAERDALN  IETAIKTKGV  DEVTIVNILT 60
   NRSNEQRQDI  AFAYQRRTKK  ELASALKSAL  SGHLETVILG  LLKTPAQYDA  SELKASMKGL 120
   GTDEDSLIEI  ICSRTNQELQ  EINRVYKEMY  KTDLEKDIIS  DTSGDFRKLM  VALAKGRRAE 180
   DGSVIDYELI  DQDARDLYDA  GVKRKGTDVP  KWISIMTERS  VCHLQKVFER  YKSYSPYDML 240
   ESIKKEVKGD  LENAFLNLVQ  CIQNKPLYFA  DRLYDSMKGK  GTRDKVLIRI  MVSRSEVDML 300
   KIRSEFKRKY  GKSLYNYIQQ  DTKGDYQKAL  LYLCGGDD 338
```

p11

```
   PSQMEHAMET  MMFTFHKFAG  DKGYLTKEDL  RVLMEKEFPG  FLENQKDPLA  VDKIMKDLDQ 60
   CRDGKVGFQS  FFSLIAGLTI  ACNDYFVVHM  KQKGK 95
```

Figure 3.12: Amino Acid Sequences of ANXA2 and p11 from *Sus scrofa domestica*. The N-terminal Ser of annexin A2 is acetylated. Sequence coverages of annexin A2 and p11 are 100% and 95.8% (underlined amino acids not covered), respectively**.**

Although SH groups of cysteines of ANXA2 and p11 were alkylated with iodoacetamide, carbamidomethylation was found to be incomplete in the case of p11.

3.2.3   Immunochemical Detection of Annexin A2 by Western Blot Analysis

Western blot analysis allows for a rapid and simultaneous characterization of a protein by both its molecular weight and its antigenicity. During early attempts of A2t purification, aliquots taken at various stages of the purification procedure were examined for the presence of A2t. Proteins separated by SDS-PAGE were transferred to nitrocellulose membranes by semi-dry blotting. Afterwards, the membrane was incubated in fixing solution and Ponceau S staining

was employed for confirming the protein transfer to the membrane. Specific detection of ANXA2 was achieved by murine monoclonal *anti-annexin II*.

Figure 3.13 depicts the Western blot of ANXA2 and its corresponding SDS polyacrylamide gel.
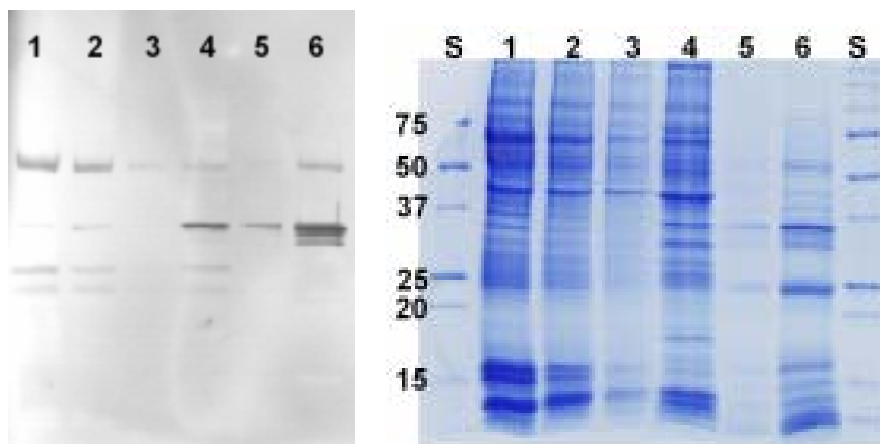
Figure 3.13: Western Blot Analysis of ANXA2. *Left hand side*: Western blot with antibodies against monomeric ANXA2. *Right hand side*: Corresponding SDS gel, Coomassie-stained. Lane 1: supernatant of washing step 1, 1:10 diluted; lane 2: supernatant of washing step 2, 1:10 diluted; lane 3: supernatant of washing step 3, 1:10 diluted; lane 4: sample after ultracentrifugation; lane 5: supernatant of DEAE batch; lane 6: supernatant of DEAE batch, 10x concentrated; S: molecular weight standard (kDa). For comparison, gel bands of the Western blot are level with the gel bands on the gel.

Lanes 1 and 2 represent supernatants after two consecutive washing steps with buffer containing $CaCl_2$ and Triton X-100, whereas no Triton X-100 was employed in washing step 3 (lane 3). After EGTA addition, the sample was ultracentrifuged and the supernatant was loaded onto the gel (lane 4). For testing the interaction of A2t with the DEAE material, the supernatant was added to the DEAE slurry; after centrifugation the supernatant was separated on the gel.

The Western blot provides evidence that ANXA2 is not present in the supernatants of the washing steps, but it appears to be present in the supernatants of the ultracentrifugation after release of A2t from the membrane by addition of EGTA as well as in the DEAE supernatants (lanes 4-6, band at ~ 35 kDa). In all samples a protein band at ~ 50 kDa was detected by the antibody against ANXA2 and two faint bands at ~ 20 and 25 kDa were as well detected. This is probably due to cross-reactions of the antibody with other proteins, or, in the latter case, with proteolytically degraded ANXA2 species.

3.2.4 Molecular Weight Determination of Intact ANXA2 and P11 by MALDI-TOFMS

Linear matrix-assisted laser desorption / ionization time-of-flight mass spectrometry (Autoflex I) of the intact, proteins revealed average molecular weights of 38.5 kDa (calculated 38.49 kDa) for ANXA2 and 10.94 kDa (calculated 10.94 kDa) for p11, respectively (**Figure 3.14**). Since intact ANXA2 proved to be challenging for MALDI-TOFMS analysis, several matrices were

employed for obtaining high-quality mass spectra. Of the matrices tested, 6-aza-2-thiothymine (ATT) yielded optimum mass spectral quality for ANXA2. Besides singly charged ANXA2 and p11 monomers, the signal of the doubly charged ANXA2 monomer was observed at ~ *m/z* 19,000. Furthermore, signals of ANXA2 / p11 adducts were observed constituting the ANXA2 / p11 heterodimer (singly and doubly charged species at ~ *m/z* 50,000 and 25,000, respectively) and the ANXA2 / (p11)$_2$ heterotrimer at ~ *m/z* 60,000.

When taking a close look at the spectrum, one can anticipate signals of the following adducts: the heterotetramer ((ANXA2 /p11)$_2$) at ~ *m/z* 100,000 as well as the ANXA2 dimer at ~ *m/z* 77,000.
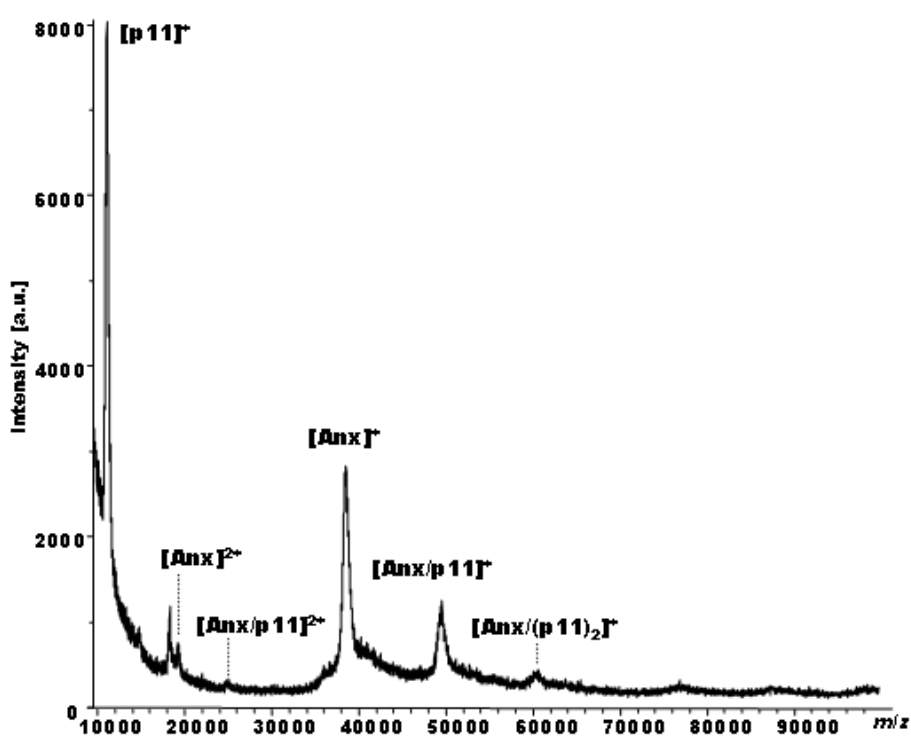
Figure 3.14: MALDI-TOF Mass Spectrum of Intact A2t Measured in Linear Positive Ionization Mode. Besides singly charged ANXA2 and p11, the spectrum contains signals of doubly charged ANXA2 as well as the following adducts: ANXA2 / p11 heterodimer (singly and doubly charged species) and ANXA2 / (p11)$_2$ heterotrimer; a.u.: arbitrary units.

3.2.5 Posttranslational Modifications

For further elucidation of the occurrence of two ANXA2 monomer bands in one-dimensional gel electrophoresis (**Figure 3.4**, lane 1, bands at 37 and 33 kDa), A2t was scrutinized for potentially present posttranslational modifications. Phosphorylation of ANXA2 amino acids S17, Y23, and S25 has been described and was shown to be capable of mediating p11-dependent translocation of ANXA2 to the cell surface [Deora, 2004].

3.2.5.1 Phosphorylation Analysis

ANXA2 and p11 were examined for phosphorylation using the Pro-Q® Diamond Phosphoprotein Gel Stain, an *in-gel* fluorescent detection assay for phosphorylation of Tyr, Ser, and Thr. In addition to the usually employed molecular weight standard for SDS-PAGE, PeppermintStick™ phosphoprotein molecular weight standard was used. It contains a mixture of two phosphorylated and four non-phosphorylated proteins, thus functioning both as molecular weight (MW) standard and as positive and negative controls. Following separation of A2t and the MW standard by gel electrophoresis, the unstained gel was subjected to the phosphostaining procedure and the stained protein bands were visualized (Figure **3.15**). On the left hand side, the positive controls appear clearly stained whereas ANXA2 is not. On the gel that is presented on the right hand side, the same samples are stained with Coomassie Blue, confirming the presence of A2t.

From the lack of phosphostaining of ANXA2 it was concluded that ANXA2 is not phosphorylated.
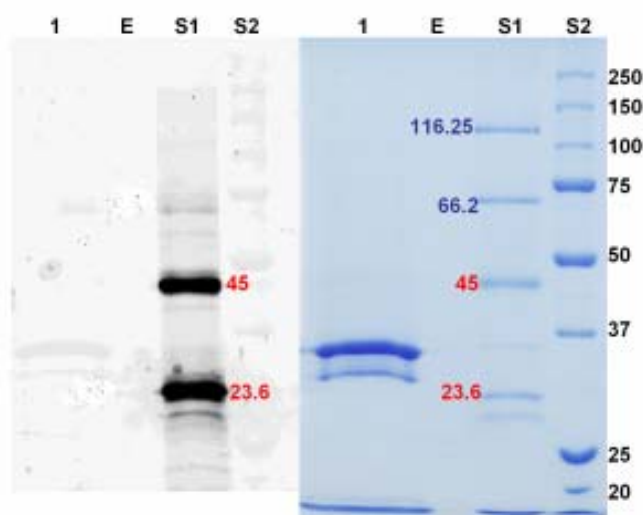
Figure 3.15: *In-Gel* Phosphorylation Analysis of A2t. Left hand side: ProQ gel staining, right hand side: Coomassie staining. Lanes 1: A2t, S1: positive and negative controls (PeppermintStick MW standard, (kDa)), S2: MW standard (kDa), E: empty lane. The contrast of the phosphostained gel was amplified.

3.2.5.2 Glycosylation Analysis

ANXA2 and p11 were scrutinized for glycosylation using the GelCode® Glycoprotein Staining Kit. Horseradish peroxidase (~ 38 kDa) and soybean trypsin inhibitor (~ 20 kDa) served as positive and negative controls, respectively. Following the gel electrophoretic separation of A2t, as well as positive and negative controls, and molecular weight standard, the gels were treated according to the manufacturer's instructions. Glycoproteins appear as magenta bands on the gel.

On the left hand side of **Figure 3.16** the gel that was prepared for glycoprotein staining is shown. Purified A2t as well as positive and negative controls were run in duplicate on the same gel and the gel was divided into two halves after gel electrophoretic separation. One half was subjected to glycoprotein staining (lanes A, B, and C) and the other half (lanes 1, 2, and 3) was stained with Coomassie Blue. The Coomassie-stained part of the gel confirms the presence of the applied proteins and their separation by gel electrophoresis.
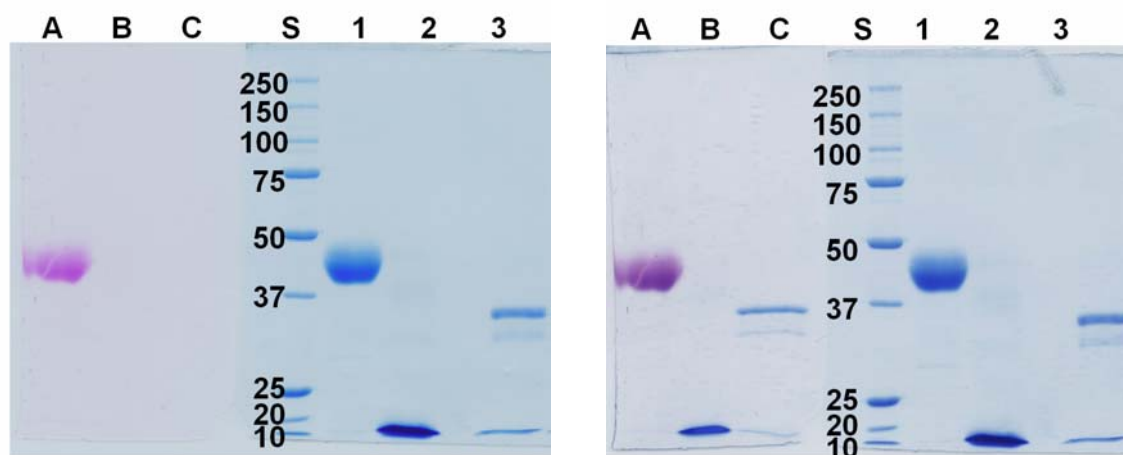
Figure 3.16: *In-Gel* Glycosylation Analysis of A2t. *Left hand side*: Lane A: positive control, lane B: negative control, lane C: A2t; subjected to gylcostaining. Lanes 1, 2, and 3, containing the same samples as lanes A, B, and C, were run as duplicates and stained with Coomassie Blue. *Right hand side*: The glyco-stained gel (lanes A, B, and C) was afterwards stained with Coomassie Blue. S: molecular weight standard (kDa).

The half that was subjected to glycostaining displays a magenta band of the positive control (horseradish peroxidase). It is clearly visible that only the positive control protein, but neither the negative control nor A2t are stained. Thus, gylcosylation of ANXA2 and p11 was excluded as posttranslational modification.

3.3 Summary

The availability of pure protein as well as an accurate and detailed description of a protein's amino acid composition and its post-translational modifications is indispensable for a successful outcome of structural studies by chemical cross-linking and mass spectrometry.

A2t was successfully purified from pig small intestinal mucosa at high purity. Purification of A2t is based on the Ca^{2+}-induced reversible association of the complex to phospholipid membranes that allows for an efficient elimination of soluble contaminants by repeated washing and centrifugation steps (**Figure 3.2**). A2t was subsequently released from the membrane by adding the chelator EGTA. A single anionic exchange chromatographic purification step on DEAE material turned out to be sufficiently effective for obtaining A2t at approximately 90% purity. A two-step NaCl gradient (0.01 and 1 M) during anionic exchange chromatography yielded A2t in the flow-through fraction, while accompanying proteins bound to the DEAE column and were released at high NaCl concentration (1 M) (**Figures 3.3** and **3.4**). Omitting the second, cationic exchange chromatography step, which did not improve A2t purity, resulted in a two day shorter purification procedure.

Overall, the applied purification protocol is simple and quick, and efficiently removes undesired proteins with minimum effort. Nevertheless, a number of challenges emerged during the purification procedure. In the beginning, too much of intestinal tissue was scraped off, which led to exceedingly impure samples. Concomitantly, the large amount of starting material made the purification laborious. In one particular case, the intestinal nematode *Ascaris suum* was observed. It is the most common endoparasite in pig with a prevalence of a few percent. The 15-40 cm long nematode leads to reduced weight gain, thus constituting an economic loss (www.animal-health-online.de/swpar/). During the latter steps of the purification procedure, protein precipitation during dialysis or after thawing of frozen samples was a major issue. It was furthermore necessary to prepare a fresh DEAE column for each use. In case a reused column was employed, A2t bound tightly to the column and could only be released with 1 M NaCl. Extensive washing of the column could not eradicate this problem.

Conclusively, for a successful purification of A2t it is crucial to scrape off only the thin mucosal layer, to avoid repeated freezing and thawing, and to employ freshly prepared DEAE columns. In total, five attempts had to be undertaken before 12 mg A2t were obtained at high purity from ~ 900 g mucosal scrapings.

Purified A2t complex was identified and characterized with respect to amino acid sequences of the ANXA2 and p11, their molecular weights, and their post-translational modifications.

The applicability of the purification protocol and selective enrichment of A2t were confirmed by Western blot analysis (**Figure 3.13**). Due to the pI value of A2t being close to 7.5, A2t was obtained in the flow-through fraction, while proteins possessing pI values well below 7.5 were bound to the anionic exchange column (**Figure 3.3**). The proteins obtained in the different fractions of the anionic exchange chromatographic separation were analyzed by peptide mass fingerprinting (**Figure 3.5**). The impurities in the flow-through fraction were identified to be C-terminally truncated ANXA2 and L36 (galectin). Other annexins, such as ANXA4 and ANXA6, were co-purified, demonstrating the selective enrichment of Ca^{2+}-binding proteins. Native gel electrophoresis under non-reducing conditions confirmed the existence of intact, non-covalent A2t, and revealed that the complex most likely exists in an oligomeric, presumably octameric, state (**Figure 3.9**). Molecular weight determination of p11 and ANXA2 monomers by MALDI-TOFMS revealed molecular weights of 38.5 and 10.94 kDa, respectively (**Figure 3.14**).

From the amino acid sequences of different vertebrate species and porcine sequences from other sources, the complete porcine amino acid sequence was deduced (**Figures 3.10** and **3.11**). The amino acid sequences of both p11 and ANXA2 were scrutinized and full sequence coverage was obtained for ANXA2 (**Figure 3.12**). The N-terminus of ANXA2 was confirmed to be acetylated. For p11 a sequence coverage of 95.8 % was obtained (**Figure 3.12**). Furthermore, it is anticipated that the faster migrating ANXA2 gel band at ~ 33 kDa (**Figure 3.4**) comprises an N-terminally truncated ANXA2 species resulting from *in-situ* proteolysis in the intestine (**Figure 3.8**).

When examining A2t for posttranslational modifications, neither phosphorylation nor glycosylation of ANXA2 were observed, as indicated by negative phosphorylation (**Figure 3.15**) and glycosylation (**Figure 3.16**) *in-gel* staining.

Altogether, A2t was successfully purified and thoroughly characterized, which is indispensable for the following structural characterization.

4   Structure Determination of the Annexin A2 / P11 Complex

The quaternary structure of A2t was investigated by chemical cross-linking, mass spectrometry, and computational protein-protein docking (Schulz *et al.*, 2007). Cross-linking reactions between ANXA2 and p11 were conducted using four amine-reactive homobifunctional cross-linking reagents with different spacer lengths. Nano-HPLC/nano-ESI-FTICRMS and MALDI-TOFMS were used for analyses of the *in-gel*-digested reaction mixtures. Increased confidence in assigning cross-linked products was achieved by employing 1:1 mixtures of non-deuterated and deuterated derivatives of cross-linking reagents that generate distinct doublet signals in the mass spectra. Furthermore, tandem mass spectrometry (MS/MS), proved to be invaluable for an exact localization of cross-linked amino acid residues and confirmation of correct cross-linked product assignment. Based on the distance restraints derived from chemical cross-linking experiments, low-resolution structural models of A2t were generated by computational docking.

4.1   Chemical Cross-Linking of A2t

The homobifunctional NHS esters sDST (disulfosuccinimidyl tartrate), BS^3-d_0/d_4 (*Bis*(sulfosuccinimidyl) suberate), BS^2G-d_0/d_4 (*Bis*(sulfosuccinimidyl) glutarate), and DSA-d_0/d_8 (disuccinimidyl adipate) (**Tables 1-1 and 1-4**) were used for chemical cross-linking as described in the *Experimental Procedures*. With the exception of DSA, the employed cross-linking reagents contain a sulfonate group at the NHS moiety, which improves water solubility. NHS esters are highly reactive towards primary amines, i. e., ε-amine groups of lysines and the free N-terminus of a protein, but they are also susceptible to hydrolysis. Upon cross-linking, NHS esters create amide bonds resulting in mass increases of 113.995 u (sDST), 138.068 u (BS^3-d_0), 96.021 u (BS^2G-d_0), and 110.037 u (DSA-d_0), respectively. Peptides, which are modified by a partially hydrolyzed cross-linker exhibit mass increases of 132.006 u (sDST), 156.079 u (BS^3-d_0), 114.032 u (BS^2G-d_0), and 128.047 u (DSA-d_0), respectively. BS^3, BS^2G, and DSA were employed as 1:1 mixtures of their non-deuterated (d_0) and deuterated (d_4 for BS^3 and BS^2G, and d_8 for DSA) species in order to facilitate the identification of cross-linked products by means of their distinct doublet isotope patterns with mass differences of 4.025 u (d_0/d_4) and 8.05 u (d_0/d_8) in the deconvoluted mass spectra. BS^3, BS^2G, and DSA were employed in 50- and 100-fold molar excess and allowed to react for 30 minutes. sDST was applied in 20- and 50-fold molar excess over the protein concentration with reaction times of 15, 30, and 60 minutes.

4.1.1 Native PAGE of Intact Cross-Linked A2t

A2t that had been cross-linked with sDST and non-cross-linked A2t were separated by native PAGE (**Figure 3.9**). Lanes 1 and 2 contain cross-linked A2t with 20- and 50-fold molar excess of cross-linker at a reaction time of 30 minutes. In lane 3, the non-cross-linked complex migrated as a broad band in the high molecular weight range. The gel bands of the cross-linked complexes appear at comparable molecular weight, however, the bands are sharper compared to those of non-cross-linked A2t. This effect is probably due to insertion of covalent linkages that prevent dissociation of the complex subunits. Most notably, the cross-linked complex migrated as a single, distinct band indicating the formation of a well-defined cross-linked complex.

4.1.2 Electrophoretic Separation of A2t Cross-Linking Reaction Mixtures Under Denaturing Conditions

The A2t cross-linking reaction mixtures were separated by one-dimensional SDS-PAGE. Complex subunits, which are not covalently connected by a cross-linker, dissociate under the denaturing conditions used for SDS-PAGE analysis.

SDS-PAGE of the sDST reaction mixtures using 20 and 50-fold molar excesses of cross-linker and reaction times between 15 and 60 minutes (**Figure 4.1 A**) revealed distinct bands of cross-linked complex subunits between ANXA2 and p11 with different stoichiometries, such as the heterodimer between ANXA2 and p11, which corresponds to a band migrating at ~ 50 kDa. Additionally ANXA2 and p11 monomers, ANXA2 dimer, $(ANXA2)_2$ / p11 heterotrimer, (ANXA2 / $p11)_2$ heterotetramer (A2t), as well as multimers were observed (**Figure 4.1 A**).

Identical patterns of SDS-PAGE gel bands were obtained when employing the isotope-labeled cross-linkers BS^3-d_0/d_4, BS^2G-d_0/d_4, and DSA-d_0/d_8 (**Figure 4.1 B**). In addition to the above mentioned cross-linked species, a p11 dimer $((p11)_2)$ was observed.

Overall, the efficiency of chemical cross-linking was rather low and seemed to be influenced by excess of reagent, but not by incubation time (**Figure 4.1 A**). Amongst the reaction times no difference in cross-linking yield was observed, whereas higher concentrations of cross-linking reagents (50-fold molar excess) led to an overall higher yield of cross-linked species. Increasing the excess of cross-linking reagent yielded higher molecular weight cross-linked species. Again, reaction times had no influence on the cross-linking patterns. The most intense gel bands were those of the monomer components ANXA2 and p11 (**Figures 4.1 A** and **B**), which dissociate during SDS-PAGE of the non-covalent complex.

A

B
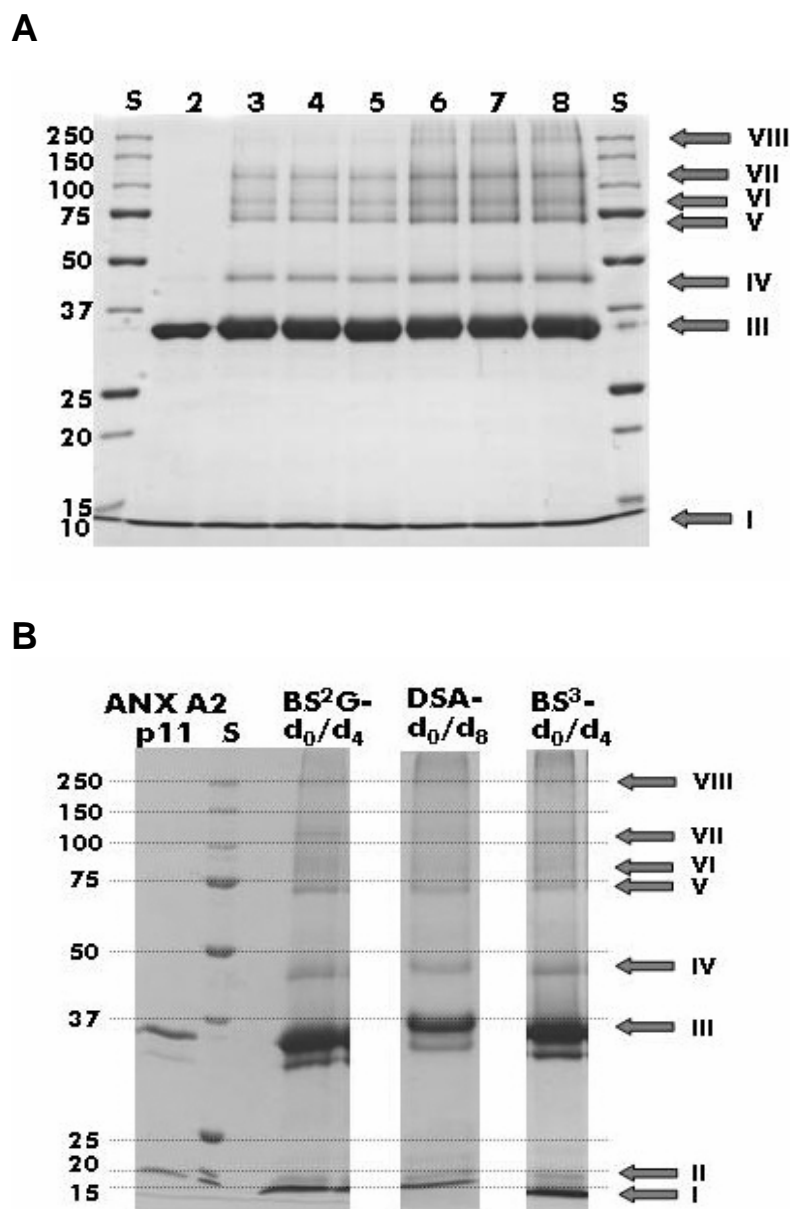
Figure 4.1: SDS-PAGE of Cross-Linked A2t. The formation of distinct cross-linked complex subunits is displayed: I: p11, II: $(p11)_2$, III: ANXA2, IV: ANXA2 / p11, V: $(ANXA2)_2$, VI: $(ANXA2)_2$ / p11, VII: (ANXA2 / p11)$_2$, VIII: (ANXA2 / p11)$_4$. 5% stacking/ 12% resolving gel. *A*, A2t cross-linked with sDST. Lane 2: non-cross-linked A2t, lanes 3-5: 20-fold excess of sDST, 15, 30, 60 min incubation time, lanes 6-8: 50-fold excess of sDST, 15, 30, 60 min incubation time. *B*, A2t cross-linked with BS^{2}G-d_0/d_4, DSA-d_0/d_8, and BS^3-d_0/d_4.; 50-fold excess of cross-linkers BS^{2}G-d_0/d_4, DSA-d_0/d_8, and BS^3-d_0/d_4, 30 min incubation time. S: molecular weight standard (kDa).

For subsequent MS analysis of the cross-linked products, the bands of interest were excised from the gel and subjected to proteolytic *in-gel* digestion. A2t cross-linked with sDST was digested with endoproteinase AspN, whereas A2t cross-linked with BS^3-d_0/d_4, BS^{2}G-d_0/d_4, and DSA-d_0/d_8 was digested with trypsin.

4.2 Identification of Cross-Linked Products by Mass Spectrometric Analysis

The peptide mixtures derived from *in-gel* proteolysis of the gel bands of cross-linked A2t subunits (**Figures 4.1 A** and **B**) were analyzed by MALDI-TOFMS and nano-HPLC / nano-ESI-FTICRMS. The obtained *m/z* values were compared to *in-silico* digests of cross-linking reaction mixtures and the masses matching unmodified peptides and cross-linked products were assigned.

4.2.1 Mass Spectrometric Analysis of A2t Cross-Linking Reaction Mixtures

The peptide mixtures of *in-gel* digested cross-linked A2t species were analyzed by nano-HPLC / nano-ESI-FTICRMS (Apex II) and MALDI-TOFMS (Autoflex I). Additionally, peptide mixtures derived from heterodimer (ANXA2 / p11, ~ 50 kDa) and heterotetramer ((ANXA2 / p11)$_2$, ~ 100 kDa) complexes were analyzed by MS/MS on a hybrid linear ion trap / FTICR mass spectrometer (LTQ-FT). Mass spectra of MALDI-TOFMS and ESI-FTICRMS (Apex II) measurements were evaluated manually. Data analysis of tandem mass spectrometric data was performed by dividing chromatogram of the total ion current (TIC) into five-minute fractions. The obtained average mass spectra were deconvoluted using the Mascot Distiller software

Signal intensities of cross-linked products are very low compared to signals of unmodified peptides and labeling of the mass spectra had to be done manually. Unfortunately, the software provided by the manufacturer and other available MS data evaluation programs do not accurately assign signals. Moreover, deconvolution performed by the XMASS software was often unsatisfactory. Scrutinizing a large complex, such as A2t, yields a high number of peptide signals further complicating cross-linked product assignment. MS analysis was also hampered by the long hydrophobic stretch in p11 that is not cleavable by commonly employed specific proteases leading to cross-linked products with long retention times and high *m/z* values.

4.2.2 Evaluation of Cross-Linking Data

The *m/z* values obtained from MALDI-TOFMS and nano-HPLC / nano-ESI-FTICRMS were assigned to peptides and cross-linked products by comparison with *in-silico* generated possible cross-linked products and peptides using the GPMAW software. Product ion mass spectra of identified cross-linked peptides were processed with the Mascot Distiller software and assigned by MS2Assign.

Application of isotope-labeled cross-linkers greatly facilitated cross-linked product identification in the mass spectra as only signals displaying the characteristic doublet are used for searching cross-linked products. This constitutes an enormous facilitation in data evaluation.

It turned out that high mass accuracies alone are no longer sufficient for a correct cross-linked product assignment when dealing with high molecular weight protein assemblies, such the A2t complex, as various candidate cross-linked peptides have molecular masses falling within the mass accuracy of the mass spectrometer. At that point, tandem mass spectrometric (MS/MS) analyses provide the means for a correct cross-linked product assignment and in case two or more potentially cross-linkable amino acids are present, allow for an exact localization of the cross-link. H_2O and NH_3 losses were frequently observed in product ion mass spectra as well as cleavages at the amide bond created between an amine group of the peptide and the cross-linking reagent.

Tandem mass spectrometric analysis of ambiguous cross-linked products revealed that C-terminal cleavage of modified lysines is impaired, if not impossible. Thus, tryptic cleavage C-terminal to a cross-linked lysine residue has to be considered less likely, yet it cannot be definitely repudiated when MS/MS data are not available. N-terminal cleavage of Asp and Glu by endoproteinase AspN is not affected in case the preceding Lys is modified.

4.3 Cross-Linked Products of A2t

4.3.1 Intermolecular Cross-Linked Products Between ANXA2 and P11

In case of investigating protein interfaces, intermolecular cross-linked products between the binding partners are the most desirable cross-linked products as they provide important information on contact sites within protein assemblies. For A2t, five intermolecular cross-linked products between ANXA2 and p11 were identified (**Table 4-1**).

Table 4-1: Intermolecular Cross-Linked Products of A2t

Modified residues printed in bold were unambiguously identified by MS/MS data. [x]: no MS/MS data available, [xx]: poor quality MS/MS data, M: oxidized Met, C: Cys is not carbamidomethylated. [2] Nano-HPLC / nano-ESI-FTICRMS (Apex II), [3] Hybrid linear ion trap / FTICRMS (LTQ-FT). Column *Sample* states the molar excess of cross-linker, incubation time, and protease (T: trypsin, AspN).

Gel Band	Sequence ANXA2	Sequence p11	Reagent	Modified Residue ANXA2	Modified Residue p11	*m/z* $[M+H]^+_{exp}$	*m/z* $[M+H]^+_{cal}$	Δ m (ppm)	Sample
ANXA2 / p11 50kDa	33-50 (F)/DAERDALNIETAIKTK GV/(D)	52-58 (V)/DKIMKDL/(D)	sDST	K46, K48	K53, K56	**2707.399**	2707.382	6.2	50x, 15 min, AspN, [2]
	109-124 (Y)/DASELKASMKGLGTD E/(D)	52-56 (V)/DKIMK/(D)	sDST	**K118**	**K56**	**2431.130**	2431.121	3.7	50x, 30 min, AspN, [3]
	165-185 (G)/DFRKLMVALAKGRRA EDGSVI/(D)	43-46 (L)/ENQK/(D)	sDST	K168, K175	K46	2963.530 2963.527 2963.501	2963.537	2.3 3.3 12.1	20x, 15 min, AspN, [2] 20x, 30 min, AspN, [2] 50x, 15 min, AspN, [2]
	168-177 (R)/KLMVALAKGR/(R)	47-62 (K)/DPLAVDKIMKD LDQCR/(D)	DSA d_0/d_8	K168, K175	K53, K56	d_0 3071.643 d_8 3079.667	3071.641 3079.691	0.6 7.7	100x, 30 min, T, [3], x
	220-230 (R)/SVCHLQKVFER/(Y)	92-95 (K)/QKGK	BS³ d_0/d_4	S220, K226	K93, K95	d_0 2000.077 d_4 2004.101	2000.075 2004.100	0.9 0.4	100x, 30 min, T, [3], xx

Figure 4.2 A shows the chromatogram of the total ion current (TIC) during nano-HPLC / nano-ESI-FTICRMS analysis of the peptide mixture from the ANXA2 / p11 heterodimer (**Figure 4.1 A**, cross-linker sDST, lane 7, band IV). In the time range between 15 and 20.5 minutes, the corresponding average mass spectrum (**Figure 4.2 A** inset and **Figure 4.2 B**), exhibits the signal of a triply charged ion of a cross-linked product (**Table 4-1**) between ANXA2 (amino acids 109-124) and p11 (amino acids 52-56) at *m/z* 811.048 (**Figure 4.2 B** inset), which was subsequently subjected to MS/MS analysis. Interestingly, the mass spectrum also contains the unmodified ANXA2 peptide comprised of amino acids 109-124 as well as the same peptide that is modified with a partially hydrolyzed cross-linker.
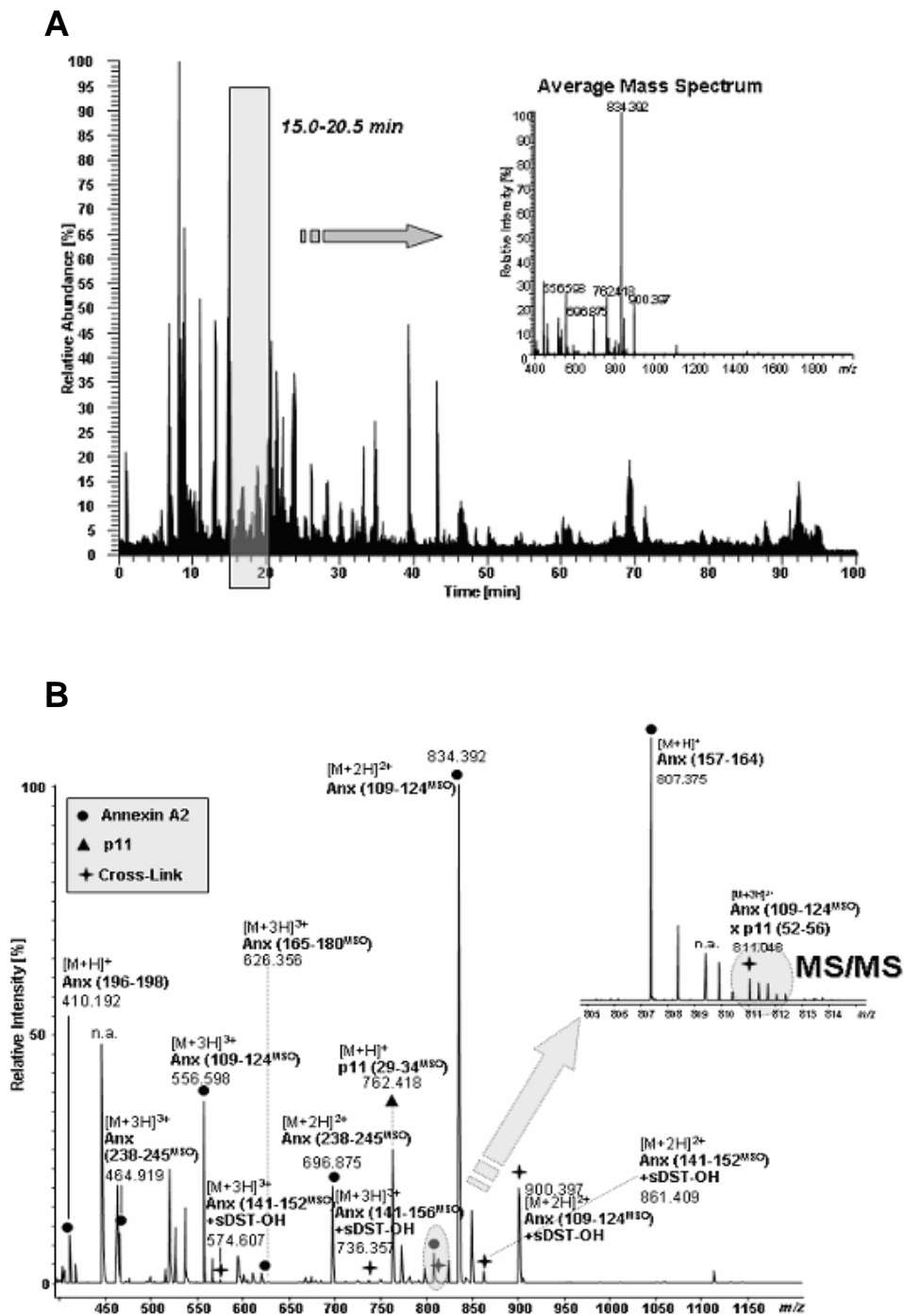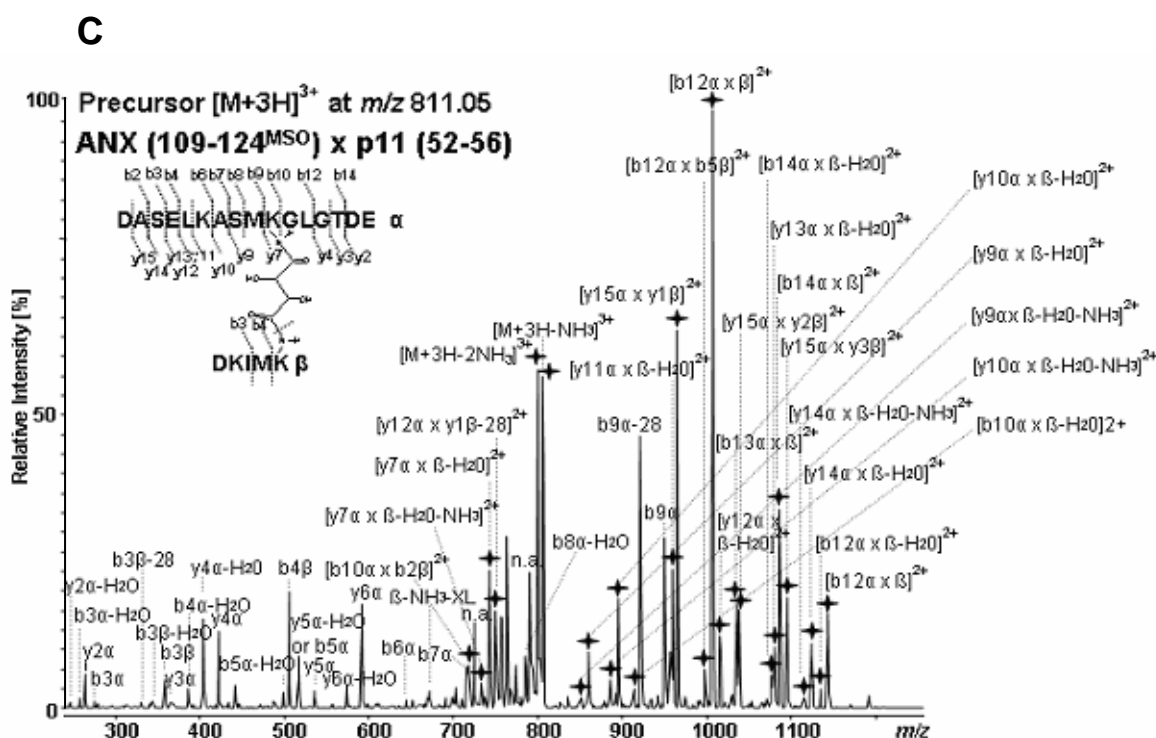
Figure 4.2: Tandem Mass Spectrometric Analysis of A2t Cross-Linked with sDST. Analysis of an AspN digest of A2t cross-linked with sDST (50-fold excess, 30 min. incubation time, heterodimer gel band IV, Figure 4.1 A). *A*, Total Ion Current. The chromatogram of the total ion current (TIC) during nano-HPLC / nano-ESI-FTICRMS analysis was divided into five-minute fractions and the obtained average mass spectra were deconvoluted. *B*, Average Mass Spectrum. Average mass spectrum from the 15-20.5 minute fraction in the TIC. The spectrum contains peptide signals of ANXA2 (circle) and p11 (triangle), as well as signals of cross-linked products (star). sDST-OH: partially hydrolyzed sDST. The inset shows the signal at *m/z* 811.048 of a triply charged intermolecular cross-linked product (Table 4-1) between ANXA2 and p11. The signal highlighted in the inset was analyzed by MS/MS. *C*, Product Ion Mass Spectrum. Product ion mass spectrum of the triply charged precursor (*m/z* 811.048) of the cross-linked product between ANXA2 and p11. The star denotes signals of cross-linked products. n.a.: not assigned, MSO: Met sulfoxide, XL: cross-linker. Cleavage of the cross-linker at the amide bond has been observed as indicated by a dotted line (corresponding signal labeled *β*-NH$_3$-XL).

MS/MS analysis of the cross-linked product between ANXA2 (amino acids 109-124) and p11 (amino acids 52-56) (**Figure 4.2 C**) led to the identification of Lys-118 in ANXA2 and Lys-56 in p11 to be cross-linked. Cleavages do not only occur at the peptide backbone, but also the amide bond connecting the peptide with the cross-linker can be cleaved. Additional H$_2$O and NH$_3$ losses of fragment ions were frequently observed.

The application of isotope-labeled reagents significantly facilitates cross-linked product identification based on the characteristic isotope patterns. By employing 1:1 mixtures of non-deuterated and deuterated cross-linker derivatives BS^3-d_0/d_4 and DSA-d_0/d_8 two intermolecular cross-linked products were obtained (**Table 4-1**). When employing DSA-d_0/d_8, amino acid sequence 168-175 in ANXA2 was found to be cross-linked to 47-62 in p11. In this cross-linked product, Cys 61 of the p11 sequence was not carbamidomethylated. Unfortunately, no MS/MS data were obtained for this cross-linked product, making it impossible to discriminate which of the lysines (Lys-168 or 175 in ANXA2 and Lys-53 or 56 in p11) are actually cross-linked.

Furthermore, a BS^3-d_0/d_4-cross-linked product was identified between amino acid sequence 220-226 in ANXA2 and the C-terminal amino acid residues 91-95 in p11. MS/MS data were obtained for this cross-linked product, but they were of minor quality (data not shown). Either Ser-220 or Lys-226 in ANXA2 is cross-linked to either Lys-93 or Lys-95 in p11.

Conclusively, the just described three cross-linked products (ANXA2 109-124 x p11 52-56, ANXA2 168-175 x p11 47-62, and ANXA2 220-226 x p11 91-95) were used as distance restraints for creating the structural models of A2t by computational protein-protein docking (chapter 4.4).

Two further cross-linked products obtained from cross-linking with sDST (**Table 4-1**), namely ANXA2 (amino acids 165-185) cross-linked with p11 (amino acids 43-46) and ANXA2 (amino acids 33-50) cross-linked with p11 (amino acids 52-58), were identified merely based on exact mass measurements, i.e., without isotope-labeled cross-linker or MS/MS data, and are therefore less reliable. For this reason, these two cross-links were not used for creating the structural model of A2t. Nevertheless, the cross-linked product between amino acid sequence 165-185 in ANXA2 and 43-46 in p11, which was identified in three different samples, encompasses the same regions as the above described cross-linked product obtained with DSA-d_0/d_8, and proved to perfectly fit the completed model structure (see chapter 4.4).

4.3.2 Intramolecular Cross-Linked Products Within ANXA2 and P11

In addition to the identified intermolecular cross-linked products, a number of intramolecular cross-linked products were identified within ANXA2 and p11 (**Table 4-2**). These can be either interpeptide, i. e., between two peptides of the same molecule, or intrapeptide, i. e., within one peptide.

For ANXA2, a large number of intramolecular cross-linked products were unambiguously assigned (**Table 4-2**), however, some of the obtained *m/z* values could not be distinctly assigned to a unique cross-linked product. In these cases, both possible cross-linked products are listed in **Table 4-2**. **Figure 4.3** depicts the MALDI-TOF mass spectrum of a tryptic digest of the ANXA2 / p11 heterodimer band cross-linked with BS^3-d_0/d_4. The inset shows the doublet signals of a cross-linked product spaced at 4 amu. However, the signal could not be unambiguously assigned to a unique cross-linked product. Two different intramolecular cross-linked products within ANXA2 are possible, but their respective masses differ only by 0.09 u (**Table 4-2,** 50 kDa).

The same cross-linked product was identified from the ANXA2 monomer band by MALDI-TOFMS and ESI-FTICRMS analyses (**Table 4-2**). But even with the higher mass accuracy provided by ESI-FTICRMS, it is still not possible to tell, which of the two potential intramolecular cross-linked products is correct.
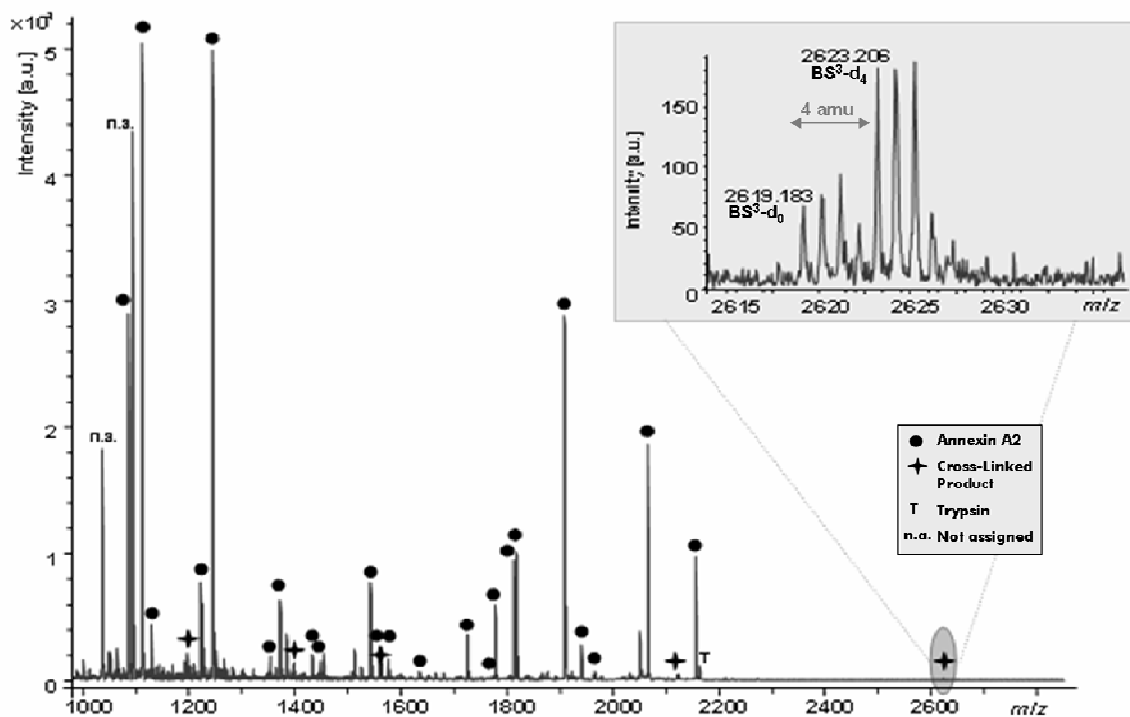
Figure 4.3: MALDI-TOF Mass Spectrum of A2t Cross-Linked with BS^3. Analysis of tryptic digest of A2t cross-linked with BS^3 (50-fold excess, 30 min. incubation time, heterodimer gel band). The inset shows the doublet isotopic pattern of an intramolecular cross-linked species within ANXA2.

Figure 4.4 shows the deconvoluted ESI-FTICR mass spectrum of the tryptically digested ANXA2 monomer band (**Figure 4.1 B**, band III) using DSA-d_0/d_8 as cross-linker. The majority of signals was assigned to ANXA2 peptides, and four signals were assigned to trypsin autolytic peptides. The spectrum contains two cross-linked products: a peptide modified by a hydrolyzed cross-linker and a cross-linked product that cannot be assigned unambiguously based on its exact mass alone. It is either an interpeptide cross-linked product between ANXA2 peptides 279-285 and 148-156 or an intrapeptide cross-linked product within ANXA2 peptide 233-248 (**Table 4-2**). Judging from the crystal structure of ANXA2, formation of the intrapeptide cross-linked product seems to be more likely than the interpeptide cross-linked product in terms of CA-CA distances of the involved lysine residues (**Table 4-3**). Furthermore, the lysine residues in the amino acid sequence 233-248 were oftentimes found to be modified by cross-linking reagent (**Table 4-2**).
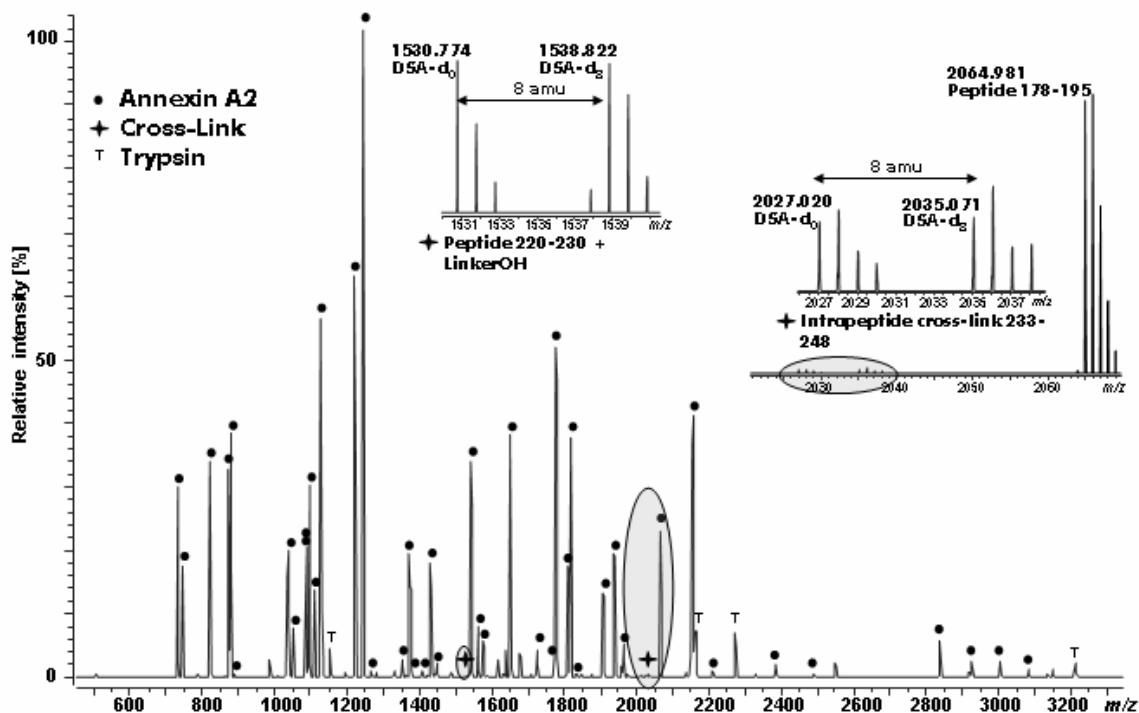
Figure 4.4: Deconvoluted Nano-ESI FTICR Mass Spectrum of A2t Cross-Linked with DSA-d_0/d_8.
Tryptic digest of A2t cross-linked with DSA-d_0/d_8, 100-fold excess, 30 min; ANXA2 monomer gel band).
Circle: ANXA2 peptides, T: trypsin autolysis products, star: cross-linked products. The insets show
signals of ANXA2 peptide 220-230, which has been modified by a partially hydrolyzed cross-linker as
well as a possible intrapeptide cross-linked product of ANXA2 (aa 233-248) with the characteristic
isotope pattern (8 u spacing). Please note that this signal could also correspond to an intramolecular
cross-linked product of ANXA2 (aa 279-285 connected with aa 148-156).

An interpeptide cross-linked product was identified between ANXA2 sequences 205-211
and 245-248, which was observed for all the employed cross-linking reagents. The product ion
spectrum (**Figure 4.5**) of the interpeptide BS[3]-d_0/d_4-cross-linked product between peptides 205-
211 and 245-248 (precursor ions shown as an inset) revealed Lys-205 and Lys-245 to be cross-
linked. The signals assigned to cross-linked products are marked with a star.
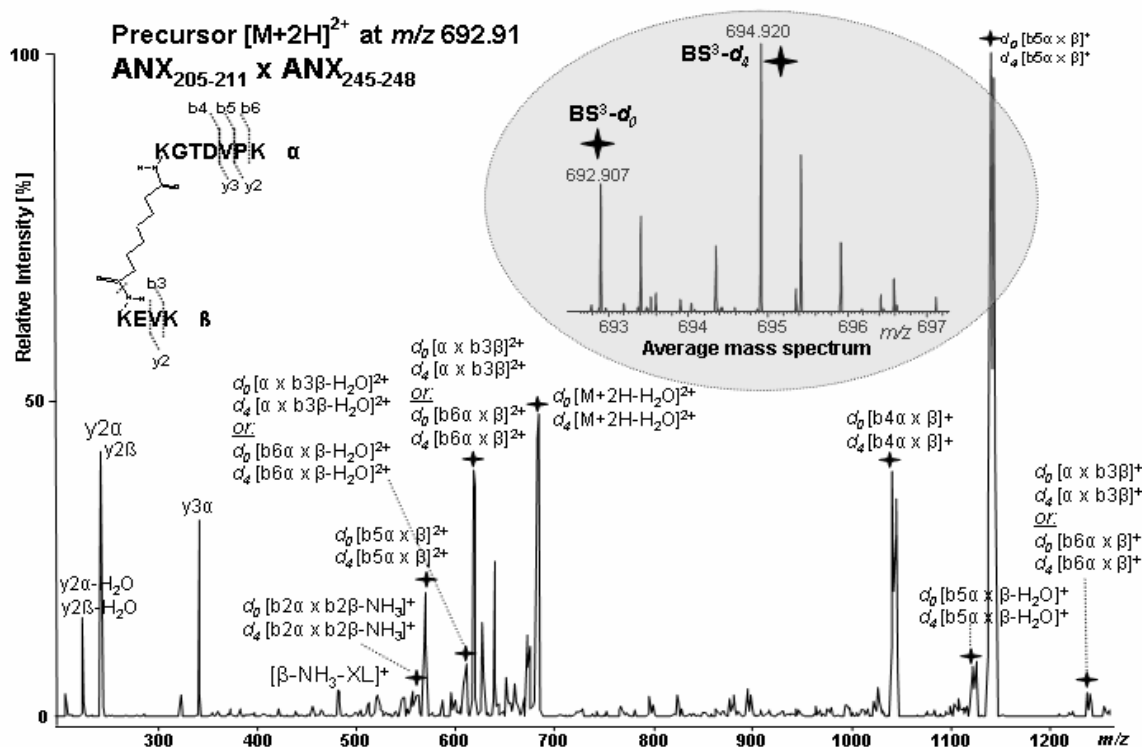
Figure 4.5: Product Ion Mass Spectrum of an Intramolecular Cross-Linked Product. Product ion spectrum of the doubly charged precursor (m/z 692.907 for d_0 species) of the cross-linked product between ANXA2 sequences 205-211 and 245-248 (BS^3-d_0/d_4, 50-fold excess, 30 min. incubation time, heterotetramer gel band, **Figure 4.2B**, band IV). Evaluation of the product ion spectrum led to the identification of K-205 and K-245 to be cross-linked. The star denotes signals of cross-linked products. Cleavage of the cross-linker at the amide bond (β-NH$_3$-XL) was also observed.

MS/MS analysis of the intrapeptide cross-linked product in ANXA2 peptide sequence 220-230 between Lys-226 and Ser-220 (calculated m/z 1512.763 for DSA-d_0 and m/z 1520.813 for DSA-d_8, **Table 4-2**) provided evidence for Ser-220 being susceptible for chemical cross-linking using NHS esters (data not shown). OH groups of serine and tyrosine residues have earlier been described to react with NHS esters (Swaim *et al.* 2004).

For peptides containing two or more lysines, there exists the possibility that the cross-linker forms an intrapeptide cross-linked product, which is subsequently cleaved by the protease between the two cross-linked residues. The created product will exhibit the same mass as a peptide that is modified by a partially hydrolyzed cross-linker. Based on MS/MS data, the cross-linked product of ANXA2 sequence 196-211 (m/z 1918.027 d_0 and 1922.054 d_4) cross-linked with BS^3-d_0/d_4 (**Table 4-2**) was identified as an interpeptide cross-linked product between K-203 and K-205 rather than the peptide modified by a partially hydrolyzed cross-linker. Please note that cross-linked amino acid residues, which were unambiguously identified by tandem mass spectrometry, are printed in bold in **Table 4-2**.

Table 4-2: Intramolecular Cross-Linked Products of A2t

Modified residues printed bold were unambiguously identified by MS/MS data. [x]: no MS/MS data available, M: oxidized Met. [1] MALDI-TOFMS (Autoflex I), [2] Nano-HPLC / nano-ESI-FTICRMS (Apex II), [3] Hybrid linear ion trap / FTICRMS (LTQ-FT). Column *Sample* states the molar excess of cross-linker, incubation time, and protease (T: trypsin, AspN).

Gel Band	Sequence ANXA2	Sequence p11	Reagent	Type	Modified Residue ANXA2	Modified Residue p11	*m/z* [M+H]$^+_{exp}$	*m/z* [M+H]$^+_{cal}$	Δ m (ppm)	Sample
~ 11kDa p11	---	32-53 (R)/VLMEKEF PGFLENQKD PLAVDK/(I)	d_0/d_8 DSA	Intrapept Xlink	---	K36 x K46 K36 x K53 K46 x K53	d_0 2673.323 d_8 2681.361	2673.348 2681.398	9.3 13.7	100x, 30 min, T, [1]
~ 37kDa ANXA2	1-12 Ac-STVHEILCKLSL/(E) **1-4 x 5-12** Ac-STVH/(E) x (H)/EILCKLSL/(E)	---	d_0/d_4 BS^3	LinkerOH / Intram Interpept	K9 / K9 x S1 K9 x T2	---	d_0 1597.750 d_4 1601.807	1597.851 1601.876	63.2 43.0	50x, 30 min, AspN, [1]
	1-4 325-336 Ac-STVH/(E) x (G)/DYQKALLYLCGG/(D)	---	sDST	Intramol Interpept	S1 x K328, T2 x K328	---	1941.892	1941.890	1.0	50x, 30 min, AspN, [2]
	69-108 (Q)/DIAFAYQRRTKKELASALKSA LSGHLETVILGLLKTPAQY/(D)	---	sDST	Intrapept Xlink	K79 x K80, K79 x K87, K79 x K103, K80 x K87, K80 x K103, K87 x K103	---	4515.527	4515.461	14.6	50x, 30 min, AspN, [2]
	141-152 (Q)/EINRVYKEMYKT/(D) **141-147 x 148-152** (Q)/EINRVYK/(E) x (K)/EMYKT/(D)	---	sDST	LinkerOH or Intramol Interpept	K147, K151 or K147 x K151	---	1705.809 1705.812	1705.810	0.5 1.1	50x, 30 min, AspN, [2]
	165-180 (G)/DFRKLMVALAKGRRAE/(D)	---	sDST	Intrapept Xlink	K168 x K175	---	1975.053 1975.065 1975.059 1975.044 1975.049 1975.051	1975.054	0.5 5.5 2.5 5.0 2.5 1.5	20x, 15 min, AspN, [2] 20x, 30min, AspN, [2] 20x, 60 min, AspN, [2] 50x, 15 min, AspN, [2] 50x, 30min, AspN, [2] 50x, 60 min, AspN, [2]
	168-177 (R)/KLMVALAKGR/(R)	---	d_0/d_4 BS^{2}G	Intrapept Xlink	K168 x K175	---	d_0 1182.717 d_4 1186.710 d_0 1182.672 d_4 1186.690	1182.703 1186.728	11.8 15.1 26.2 32.0	50x, 30 min, T, [1] 100x, 30 min, T, [1]
	196-211 (R)/DLYDAGVKRKGTDVPK/(W) **196-204 x 205-211** (R)/DLYDAGVKR/(K) x (R)/KGTDVPK/(W)	---	d_0/d_8 DSA	LinkerOH / Intrapept Xlink	K203, K205, K211 / K203xK205, K203xK211, K205xK211	---	d_0 1889.929 d_8 1897.996	1889.997 1898.047	35.9 26.8	100x, 30 min, T, [1]
	220-230 (R)/SVCHLQKVFER/(Y)	---	d_0/d_4 BS^3	Intrapept Xlink	S220 x K226	---	d_0 1540.797 d_4 1544.824	1540.794 1544.819	1.9 3.2	100x, 30 min, T, [2]
	238-249 (Y)/DMLESIKKEVKG/(D) **238-245 x 246-249** (Y)/DMLESIKK/(E) x (K)/EVKG/(D)	---	sDST	LinkerOH / Intramol Interpept	K244, K245, K248 / K244xK245, K244 x K248	---	1508.754 1508.756 1508.758	1508.751	1.9 3.3 4.6	20x, 15 min, AspN, [2] 50x, 30 min, AspN, [2] 50x, 60 min, AspN, [2]
	273-283 (R)/LYDSMKGKGTR/(D) or **279-285 x 309-312** (K)/GKGTRDK/(V) x (R)/KYGK/(S)	---	d_0/d_4 BS^3	Intrapept Xlink or Intramol Interpept	K278 x K280 or K280 x K309, K280 x K312, K285 x K309, K285 x K312	---	d_0 1393.716 d_4 1397.714 d_0 1393.719 d_4 n.a.	1393.714 1397.739 or 1393.780 1397.805	1.4 16.4 or 45.9 65.1 or 3.5 n.a. or 43.7 n.a.	100x, 30 min, T, [1] 50x, 30 min, T, [1]
	279-285 x 148-156 (K)/GKGTRDK/(V) x (K)/EMYKTDLEK/(D) or 233-248 (K)/SYSPYDMLESIKKEVK/(G)	---	d_0/d_8 DSA	Intramol Interpept or Intrapept Xlink	K280 x K151, K280 x K156, K285 x K151, K285 x K156 or K244 x K245, K244 x K248, K245 x K248	---	d_0 2027.020 d_8 2035.071	2027.011 2035.061 or 2027.004 2035.054	4.4 4.9 or 7.8 8.3	100x, 30 min, T, [2]
	279-283 x 313-328 (K)/GKGTR/(D) x (K)/SLYNYIQQDTKGDYQK/(A) or 304-309 x 324-338 (R)/SEFKRK/(Y) x (K)/GDYQKALLYLCGGDD	---	d_0/d_4 BS^3	Intramol Interpept / Intramol Interpept / Intramol Interpept	K280 x K323 K280 x K328 / K307 x K328 K309 x K328	---	d_0 2619.302 d_4 2623.328 d_0 2619.304 d_4 2623.331 d_0 2619.260 d_4 2623.274	2619.305 2623.330 or 2619.276 2623.301	1.3 0.7 or 9.9 10.2 0.3 0.3 or 10.6 11.4 17.1 21.3 or 6.1 10.2	100x, 30 min, T, [2] 100x, 30 min, T, [1] 50x, 30 min, T, [1]
	304-309 x 324-338 (R)/SEFKRK/(Y) x (K)/GDYQKALLYLCGGDD	---	d_0/d_4 BS^{2}G	Intramol Interpept	K307 x K328 K309 x K328	---	d_0 2577.126 d_4 2581.145 d_0 2576.973 d_4 2581.002	2577.229 2581.254	39.9 42.2 99.3 97.6	50x, 30 min, T, [1] 100x, 30 min, T, [1]

Group	Peptide / Sequence		Crosslinker	Xlink type	Sites		Observed	Calculated	Error	Conditions
~ 50kDa ANXA2 / p11	69-108 (Q)/DIAFAYQRRTKKELASALKSA LSGHLETVILGLLKTPAQY/(D)	---	sDST	Intrapept Xlink	K79 x K80, K79 x K87, K79 x K103, K80 x K87, K80 x K103, K87 x K103	---	4515.538 / 4515.490 / 4515.490 / 4515.489	4515.461	17.0 / 6.4 / 6.4 / 6.2	20x, 30 min, AspN, 2) / 20x, 60 min, AspN, 2) / 50x, 15 min, AspN, 2) / 50x, 60 min, AspN, 2)
	77-87 (R)/RTKKELASALK/(S)	---	d_0/d_4 BS^{2}G	Intrapept Xlink	K79 x K80	---	d_0 1340.791 / d_4 1344.817	1340.790 / 1344.815	0.7 / 1.4	50x, 30 min, T, 3)
	77-87 (R)/RTKKELASALK/(S)	---	d_0/d_8 DSA	Intrapept Xlink	K79 x K80	---	d_0 1354.807 / d_8 1362.857	1354.805 / 1362.855	1.4 / 1.4	100x, 30 min, T, 3)
	77-87 (R)/RTKKELASALK/(S)	---	d_0/d_4 BS^3	Intrapept Xlink	K79 x K80	---	d_0 1382.837 / d_4 1386.862	1382.837 / 1386.862	0 / 0	100x, 30 min, T, 3)
	78-87 (R)/TKKELASALK/(S)	---	d_0/d_4 BS^{2}G	LinkerOH or Intramol interpept	K79, (K80)	---	d_0 1202.699 / d_4 1206.724	1202.699 / 1206.724	0 / 0	50x, 30 min, T, 3)
	78-87 (R)/TKKELASALK/(S)	---	d_0/d_8 DSA	Intrapept Xlink	K79 x K80	---	d_0 1198.705 / d_8 1206.757	1198.704 / 1206.754	0.8 / 2.4	100x, 30 min, T, 3)
	81-108 (K)/ELASALKSALSGHLETVILGLL KTPAQY/(D)	---	sDST	Intrapept Xlink	K87 x K103	---	3037.586 / 3037.602 / 3037.609	3037.646	19.7 / 14.4 / 12.1	50x, 15 min, AspN, 2) / 50x, 30 min, AspN, 2) / 50x, 60 min, AspN, 2)
	141-147 x 165-180 (Q)/EINRVYK/(E) x (G)/DFRKL MVALAKGRRAE/(D)	---	sDST	Intramol Interpept	K147 x K168 / K147 x K175	---	2911.545 / 2911.522 / 2911.556 / 2911.568	2911.557	4.1 / 12.0 / 0.3 / 3.7	20x, 30 min, AspN, 2) / 20x, 60 min, AspN, 2) / 50x, 30 min, AspN, 2) / 50x, 60 min, AspN, 2)
	168-177 (R)/KLMVALAKGR/(R)	---	d_0/d_4 BS^{2}G	Intrapept Xlink	K168 x K175	---	d_0 1182.648 / d_4 1186.716	1182.703 / 1186.728	46.5 / 10.1	100x, 30 min, T, 1)
	168-177 (R)/KLMVALAKGR/(R)	---	d_0/d_4 BS^{2}G	Intrapept Xlink	K168 x K175	---	d_0 1198.700 / d_4 1202.722	1198.698 / 1202.723	1.6 / 0.8	50x, 30 min, T, 3)
	168-177 (R)/KLMVALAKGR/(R)	---	d_0/d_8 DSA	Intrapept Xlink	K168 x K175	---	d_0 1196.719 / d_8 1204.769	1196.718 / 1204.768	0.8 / 0.8	100x, 30 min, T, 3)
	196-204 x 245-248 (R)/DLYDAGVKR/(K) x (K)/KEVK/(G)	---	d_0/d_4 BS^3	Intramol Xlink	K203 x K245	---	d_0 1676.925 / d_4 1680.949	1676.922 / 1680.947	1.7 / 1.1	100x, 30 min, T, 3)
	196-211 (R)/DLYDAGVKRKGTDVPK/(W)	---	d_0/d_8 DSA	LinkerOH or Intramol interpept	K203, K205	---	d_0 1889.997 / d_8 1898.046	1889.997 / 1898.047	0 / 0.5	100x, 30 min, T, 3)
	196-211 (R)/DLYDAGVKRKGTDVPK/(W)	---	d_0/d_4 BS^3	LinkerOH or Intramol interpept	K203 x K205	---	d_0 1918.027 / d_4 1922.054	1918.028 / 1922.053	0.5 / 0.5	100x, 30 min, T, 3)
	205-211 x 245-248 (R)/KGTDVPK/(W) x (K)/KEVK/(G)	---	d_0/d_8 DSA	Intramol Xlink	K205 x K245	---	d_0 1356.774 / d_8 1364.823	1356.773 / 1364.823	0.7 / 0	100x, 30 min, T, 3)
	205-211 x 245-248 (R)/KGTDVPK/(W) x (K)/KEVK/(G)	---	d_0/d_4 BS^{2}G	Intramol Xlink	K205 x K245	---	d_0 1342.758 / d_4 1346.785	1342.758 / 1346.783	0 / 1.4	50x, 30 min, T, 3)
	220-230 (R)/SVCHLQKVFER/Y	---	d_0/d_8 DSA	Intrapept Xlink	S220 x K226	---	d_0 1512.766 / d_8 1520.816	1512.763 / 1520.813	1.9 / 1.9	100x, 30 min, T, 3)
	275-304 (Y)/DSMKGKGTRDKVLIRIMVSRS EVDMLKIRS/(E)	---	sDST	Intrapept Xlink, 1xMSO	K278 x 280, K278 x K285, K278 x K301, K280 x K285, K280 x K301, K285 x K301	---	3578.924	3578.867	15.9	20x, 15 min, AspN, 2
	275-304 (Y)/DSMKGKGTRDKVLIRIMVSRS EVDMLKIRS/(E)	---	sDST	Intrapept Xlink, all MSO	K278 x 280, K278 x K285, K278 x K301, K280 x K285, K280 x K301, K285 x K301	---	3610.920	3610.871	13.5	20x, 15 min, AspN, 2
	279-283 x 313-328 (K)/GKGTR/(D) x (K)/SLYNYIQQDTKGDYQK/(A) or 304-309 x 324-338 (R)/SEFKRK/(Y) x (K)/GDYQKALLYLCGGDD	---	d_0/d_4 BS^3	Intramol Interpept or Intramol Interpept	K280 x K323, K280 x K328 or K307 x K328, K309 x K328	---	2619.183 / 2623.206	2619.305 / 2623.330 / 2619.276 / 2623.301	46.5 / 47.2 / 35.5 / 36.2	100x, 30 min, T, 1)
	304-309 x 324-338 (R)/SEFKRK/(Y) x (K)/GDYQKALLYLCGGDD	---	d_0/d_4 BS^{2}G	Intramol Interpept	K307 x K328 / K309 x K328	---	d_0 2577.160 / d_4 2581.165	2577.229 / 2581.254	26.7 / 34.4	100x, 30 min, T, 1)
~ 75kDa (ANXA2)$_2$	168-177 (R)/KLMVALAKGR/(R)	---	d_0/d_4 BS^{2}G	Intrapept Xlink	K168 x K175	---	d_0 1182.696 / d_4 1186.735 / d_0 1182.678 / d_4 1186.704	1182.703 / 1186.728	5.9 / 5.8 / 21.1 / 41.2	50x, 30 min, T, 1) / 100x, 30 min, T, 1)
	304-309 x 324-338 (R)/SEFKRK/(Y) x (K)/GDYQKALLYLCGGDD	---	d_0/d_4 BS^{2}G	Intramol Interpept	K307 x K328 / K309 x K328	---	d_0 2577.254 / d_4 2581.257 / d_0 2577.164 / d_4 2581.189	2577.229 / 2581.254	9.7 / 11.0 / 25.2 / 25.1	50x, 30 min, T, 1) / 100x, 30 min, T, 1)
~ 85kDa (ANXA2)$_2$/p11	69-108 (Q)/DIAFAYQRRTKKELASALKSA LSGHLETVILGLLKTPAQY/(D)	---	sDST	Intrapept Xlink	K79 x K80, K79 x K87, K79 x K103, K80 x K87, K80 x K103, K87 x K103	---	4515.461 / 4515.478	4515.461	0 / 3.7	20x, 15 min, AspN, 2) / 20x, 60 min, AspN, 2)
	81-108 (K)/ELASALKSALSGHLETVILGLL KTPAQY/(D)	---	sDST	Intrapept Xlink	K87 x K103	---	3037.598 / 3037.598 / 3037.591	3037.646	15.8 / 15.8 / 18.1	20x, 15 min, AspN, 2) / 20x, 30 min, AspN, 2) / 20x, 60 min, AspN, 2)
	141-147 x 165-180 (Q)/EINRVYK/(E) x (G)/DFRKL MVALAKGRRAE/(D)	---	sDST	Intramol Interpept	K147 x K168 / K147 x K175	---	2911.550 / 2911.551	2911.557	2.4 / 2.0	20x, 15 min, AspN, 2) / 20x, 30 min, AspN, 2)
~100kDa (ANXA2 /p11)$_2$	69-108 (Q)/DIAFAYQRRTKKELASALKSA LSGHLETVILGLLKTPAQY/(D)	---	sDST	Intrapept Xlink	K79 x K80, K79 x K87, K79 x K103, K80 x K87, K80 x K103, K87 x K103	---	4515.479 / 4515.482	4515.461	3.9 / 4.6	20x, 30 min, AspN, 2) / 20x, 60 min, AspN, 2)
	77-87 (R)/RTKKELASALK/(S)	---	d_0/d_4 BS^3	Intrapept Xlink	K79 x K80	---	d_0 1382.837 / d_4 1386.863	1382.837 / 1386.862	0 / 0.7	50x, 30 min, T, 3)
	78-87 (R)/TKKELASALK/(S)	---	d_0/d_8 DSA	Intrapept Xlink	K79 x K80	---	d_0 1198.708 / d_8 1206.757	1198.704 / 1206.754	3.3 / 2.4	100x, 30 min, T, 3)

168-177 (R)/KLMVALAKGR/(R)	---	d0/d8 DSA	Intrapept Xlink	K168 x K175	---	d0 1196.720 d8 1204.770	1196.718 1204.768	1.6 1.6	100x, 30 min, T, [3]
168-177 (R)/KLMVALAKGR/(R)	---	d0/d8 DSA	Intrapept Xlink	K168 x K175	---	d0 1212.716 d8 1220.766	1212.713 1220.763	2.4 2.4	100x, 30 min, T, [3]
168-177 (R)/KLMVALAKGR/(R)	---	d0/d4 BS[3]	Intrapept Xlink	K168 x K175	---	d0 1224.752 d4 1228.778	1224.750 1228.775	1.6 2.4	50x, 30 min, T, [3]
196-211 (R)/DLYDAGVKRKGTDVPK/(W) or (R)/DLYDAGVKR/(K) x (R)/KGTDVPK/(W)	---	d0/d4 BS[3]	LinkerOH or **Intramol interpept**	**K203 x K205**	---	d0 1918.029 d4 1922.050	1918.028 1922.053	0.5 1.5	50x, 30 min, T, [3]
196-211 (R)/DLYDAGVKRKGTDVPK/(W)	---	d0/d8 DSA	LinkerOH or Intramol interpept	K203, K205, K211	---	d0 1889.991 d8 1899.042	1889.997 1899.047	3.1 2.6	100x, 30 min, T, [3], [x]
205-211 x 245-248 (R)/KGTDVPK/(W) x (K)/KEVK/(G)	---	d0/d8 DSA	Intramol Xlink	K205 x K245	---	d0 1356.775 d8 1364.823	1356.773 1364.823	1.4 1.4	100x, 30 min, T, [3]
205-211 x 245-248 (R)/KGTDVPK/(W) x (K)/KEVK/(G)	---	d0/d4 BS[3]	Intramol Xlink	K205 x K245	---	d0 1384.807 d4 1388.832	1384.805 1388.830	1.4 1.4	50x, 30 min, T, [3]
220-230 (R)/SVCHLQKVFER/(Y)	---	d0/d8 DSA	Intrapept Xlink	S220 x K226	---	d0 1512.769 d8 1520.816	1512.763 1520.813	3.9 1.9	100x, 30 min, T, [3]
304-309 (R)/SEFKRK/(Y)	---	d0/d8 DSA	Intrapept Xlink	K307 x K309	---	d0 904.442 d8 912.488	904.489 912.539	51.9 55.8	100x, 30 min, T, [1]

All intramolecular cross-linked products within ANXA2 that were unambiguously assigned (**Table 4-2**) are in agreement with the crystal structure of ANXA2 (**Figure 4.6, Table 4-3**). The side chains involved in intramolecular cross-linking are displayed in blue (unambiguous) and gray (ambiguous) spheres, and the NZ and OG atoms of lysine and serine, respectively, are shown in white. The cross-linked amino acid residues are marked and NZ-NZ (or NZ-OG) distances are given.
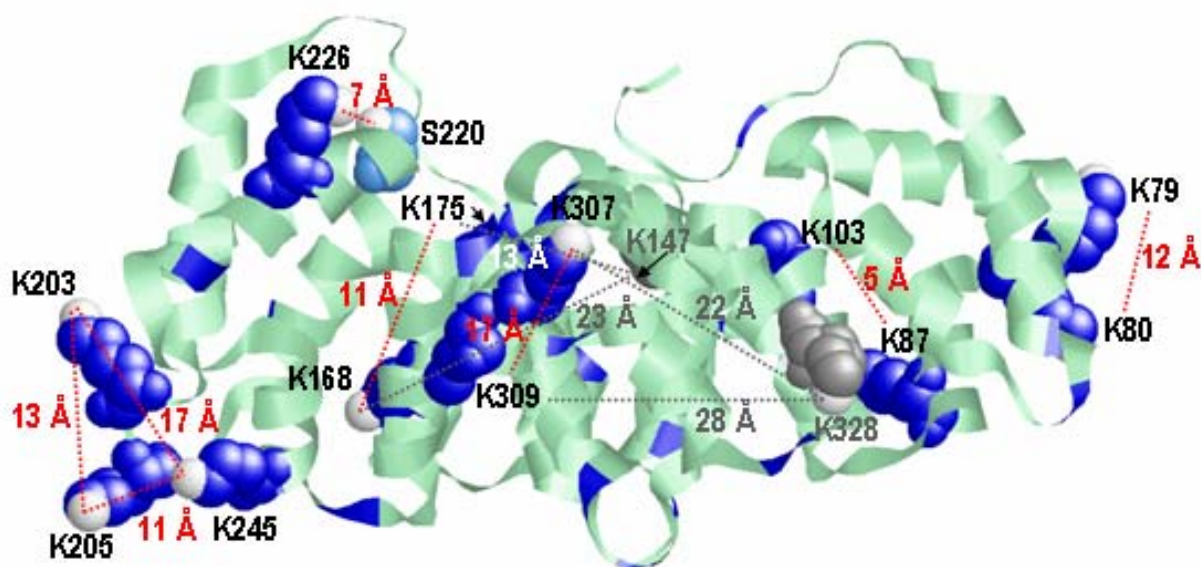
Figure 4.6: Intramolecular Cross-Linked Products within ANXA2 (PDB entry 1W7B) Visualized with RASMOL. The amino acids involved in intramolecular cross-linking are displayed and the NZ-NZ (and NZ-OG in case of K-226 x S-220) distances of the cross-linked products are indicated (red: unambiguous cross-linked products; gray: ambiguous cross-linked products) and they are in agreement with the cross-linker spacer lengths. Blue spheres: unambiguously identified lysines, gray spheres: ambiguously identified lysines; light blue: Ser-220 (unambiguous). NZ and OG atoms are shown in white.

Table 4-3 lists the amino acids involved in intramolecular cross-linking of ANXA2 and gives their respective NZ-NZ and CA-CA distances defined in the available crystal structures for monomeric ANXA2 (PDB entries 1W7B and 1XJL). Please note that lysine side chains are highly flexible, which is not reflected by crystal structure data. Given that the cross-linker spacer lengths range from 6.4 to 11.4 Å it becomes evident that the intramolecular cross-linked products match the crystal structure data.

Table 4-3: CA-CA and NZ-NZ Distances in Crystal Structures of ANXA2.
Distances printed bold constitute unambiguously assigned intramolecular cross-linked products (labeled red in **Figure 4.7**). In PDB entries 1W7B and 1XJL (both human) amino acid residue 244 constitutes an Arg rather than Lys (see **Figure 3.11**). Instead of the NZ atom the CZ atom of Arg-244 was used for measuring the distance.

| | Distance [Å] | | | |
| | PDB entry 1W7B | | PDB entry 1XJL | |
Cross-Linked Product	NZ	CA	NZ	CA
K79 x K80	**11.96**	**3.82**	**4.93**	**3.81**
K79 x K87	21.66	15.63	19.77	15.91
K79 x K103	23.72	19.16	22.90	18.98
K80 x K87	15.10	11.82	16.57	11.86
K80 x K103	18.86	16.48	20.00	16.54
K87 x K103	**5.15**	**12.92**	**8.91**	**12.93**
K147 x K151	10.55	4.89	9.78	4.49
K147 x K168	23.48	20.30	23.53	20.07
K147 x K175	12.77	13.13	12.27	13.57
K168 x K175	**11.19**	**11.02**	**13.96**	**11.03**
K203 x K205	**13.30**	**7.31**	**16.27**	**7.10**
K203 x K211	13.32	13.50	12.62	13.81
K205 x K211	21.37	14.69	20.92	14.61
K203 x K245	**16.55**	**11.44**	**17.57**	**11.50**
K205 x K245	**11.17**	**11.22**	**14.55**	**11.23**
S220 x K226	**7.03**	**19.10**	**7.54**	**10.29**
K244 x K 245	16.38	6.89	17.20 (CZ R244)	6.89
K244 x K 248	11.18	3.81	10.36	3.80
K245 x K 248	15.47	8.07	15.80	7.77
K280 x K309	25.08	20.91	25.10	20.68
K280 x K312	24.75	18.31	23.36	18.31
K285 x K309	20.96	22.78	24.41	20.09
K285 x K312	28.85	22.29	30.51	22.19
K280 x K151	43.86	39.33	43.62	39.58
K280 x K156	44.6	37.3	43.17	37.01
K280 x K156	25.12	24.27	22.41	24.57
K280 x K328	21.77	19.71	23.00	20.31
K280 x K323	13.08	13.62	15.07	14.44
K307 x K328	21.72	19.76	20.55	19.55
K309 x K328	27.55	23.95	27.90	23.85
K307 x K309	**16.68**	**5.44**	**15.96**	**5.57**
K278 x K280	16.23	7.08	16.61	7.13
K278 x K285	16.28	14.36	18.91	14.51
K278 x K301	25.69	20.34	24.37	20.11
K280 x K285	18.69	14.83	18.66	14.81
K280 x K301	37.94	26.54	37.71	26.44
K285 x K301	27.18	21.21	30.05	21.29

Only a single intramolecular cross-linked product was identified within p11, namely an intrapeptide cross-link comprising amino acids 32-53, when using DSA-d_0/d_8. Due to the lack of MS/MS data, it was impossible to sort out, which two of the three lysine residues within this p11 peptide were actually cross-linked. The three lysines in p11, which are potentially involved in

cross-linking, are displayed as gray spheres and the distances between NZ atoms are given (**Figure 4.7**).
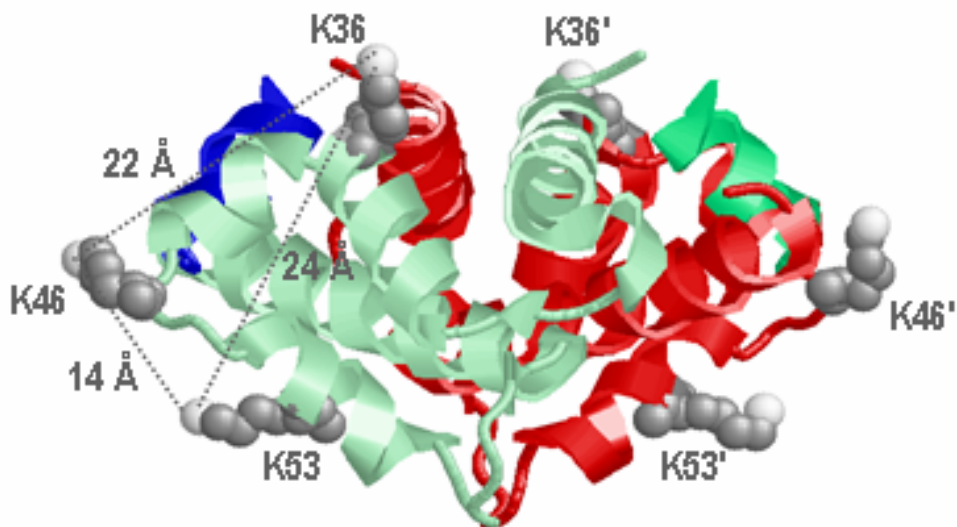
Figure 4.7: Cross-Linked Products of P11 Visualized with RASMOL. Intramolecular cross-linked products and lysines with a partially hydrolyzed cross-linker are displayed in the p11 dimer / ANXA2 peptide complex (PDB entry 1BT6, ANXA2 peptides: blue and dark green). The three lysines that are potentially modified by a partially hydrolyzed cross-linker also constitute the lysines possibly involved in intramolecular cross-linking. The distances between the NZ atoms are given. Gray spheres: ambiguously identified lysines, NZ atoms are marked white.

The fact that the intramolecular cross-links fit the crystal structures both within ANXA2 and p11 so well underlines that chemical cross-linking yields well-defined products and does not just randomly connect functional groups of amino acids.

4.3.3 ANXA2 and P11 Peptides Modified by Hydrolyzed Cross-Linker

In addition to intermolecular and intramolecular cross-linked products, a vast number of peptides were found to be modified by a cross-linker that had reacted on one side with a lysine ε-amine group or the free N-terminus, but was hydrolyzed at its second NHS moiety (**Table 4-4**). Of the 33 lysines in ANXA2, 14 lysine residues were unambiguously identified to be modified by a hydrolyzed cross-linker, either by MS/MS data or due to the fact of only one lysine (or serine) being present in the respective peptide sequence. Additionally, either K-114 or K-118 is modified by a hydrolyzed cross-linker, as well as ten further lysine residues, for which assignment is less certain. For p11, only three lysines (K-36, K-46, K-53, **Figure 4.8**) out of eleven lysines were identified to be potentially modified by a partially hydrolyzed cross-linker.

Table 4-4: Peptides Modified by Partially Hydrolyzed Cross-Linker.
Modified residues printed in bold were unambiguously identified by MS/MS data. [x)]: no MS/MS data available. [1)] MALDI-TOFMS (Autoflex I), [2)] Nano-HPLC / nano-ESI-FTICRMS (Apex II), [3)] Hybrid linear ion trap / FTICRMS (LTQ-FT). Column Sample states the molar excess of cross-linker, incubation time, and protease (T: trypsin, AspN).

Gel Band	Sequence ANXA2	Sequence p11	Reagent	Type	Modified Residue ANXA2	Modified Residue p11	m/z [M+H]$^+_{exp}$	m/z [M+H]$^+_{cal}$	Δm (ppm)	Sample
~11kD p11a	---	32-53 R/VLMEKEFPGFLENQKDPLAVDK/I	d_0/d_8 DSA	LinkerOH	---	K36, K46, K53	d_0 2691.291 d_8 2699.346	2691.359 2699.409	25.2 23.3	100x, 30 min, T, [1)]
	---	37-53 K/EFPGFLENQKDPLAVDK/I	d_0/d_8 DSA	LinkerOH	---	K46, K53	d_0 2075.037 d_8 2083.083	2075.033 2083.083	1.9 0	100x, 30 min, T, [2)]
~37kDa ANXA2	1-12 Ac-STVHEILCKLSL/(E) **1-4 x 5-12** Ac-STVH/(E) x (H)/EILCKLSL/(E)	---	d_0/d_4 BS³	LinkerOH or Intramol Interpept	K9 K9 x S1 K9 x T2	---	d_0 1597.750 d_4 1601.807	1597.851 1601.876	63.2 43.0	50x, 30 min, AspN, [1)]
	10-27 (K)/LSLEGDHSTPASAYGSVK/(A)	---	d_0/d_4 BS³	LinkerOH	K27	---	d_0 1974.973 d_4 1978.989	1974.965 1978.990	4.0 0.5	50x, 30 min, T, [1)]
	28-48 (K)/AYTNFDAERDALNIETAIKTK/(G)	---	d_0/d_8 DSA	LinkerOH	K46, K48	---	d_0 2512.218 d_8 2520.266	2512.257	15.5	100x, 30 min, T, [1)]
	109-124 (Y)/DASELKASMKGLGTDE/(D)	---	sDST	LinkerOH	K114, K118	---	1783.792 1783.803 1783.785 1783.776 1783.784 1783.791	1783.790	1.1 7.2 2.8 7.8 3.3 0.5	20x, 15 min, AspN, [2)] 20x, 30min, AspN, [2)] 20x, 60 min, AspN, [2)] 50x, 15 min, AspN, [2)] 50x, 30min, AspN, [2)] 50x, 60 min, AspN, [2)]
	141-152 (Q)/EINRVYKEMYKT/(D) **141-147 x 148-152** (Q)/EINRVYK/(E) x (K)/EMYKT/(D)	---	sDST	LinkerOH or Intramol Interpept	K147, K151 K147 x K151	--- ---	1705.809 1705.812	1705.810	0.5 1.1	50x, 30min, AspN, [2)]
	165-180 (G)/DFRKLMVALAKGRRAE/(D)	---	sDST	LinkerOH	K168, K175	---	1993.062 1993.077 1993.067 1993.060 1993.065 1993.065	1993.065	1.5 6.0 1.0 2.5 0 0	20x, 15 min, AspN, [2)] 20x, 30min, AspN, [2)] 20x, 60 min, AspN, [2)] 50x, 15 min, AspN, [2)] 50x, 30min, AspN, [2)] 50x, 60 min, AspN, [2)]
	168-177 (R)/KLMVALAKGR/(R)	---	d_0/d_8 DSA	LinkerOH	K168, K175	---	d_0 1196.740 d_8 1204.751	1196.718 1204.768	18.3 14.1	100x, 30 min, T, [1)]
	196-204 (R)/DLYDAGVKR/(K)	---	d_0/d_4 BS²G	LinkerOH	K203	---	d_0 1150.555 d_4 1154.584	1150.574 1154.599	16.5 12.9	100x, 30 min, T, [1)]
	196-204 (R)/DLYDAGVKR/(K)	---	d_0/d_4 BS³	LinkerOH	K203	---	d_0 1192.598 d_4 1196.634 d_0 1192.584 d_4 1196.606	1192.621 1196.646	19.2 31.0	100x, 30 min, T, [1)] 50x, 30 min, T, [1)]
	196-211 (R)/DLYDAGVKRKGTDVPK/(W) **196-204 x 205-211** (R)/DLYDAGVKR/(K) x (R)/KGTDVPK/(W)	---	d_0/d_8 DSA	LinkerOH or Intramol Interpept	K203, K205, K211 K203xK205, K203xK211, K205xK211	---	d_0 1889.929 d_8 1897.996	1889.997 1898.047	35.9 26.8	100x, 30 min, T, [1)]
	220-230 (R)/SVCHLQKVFER/(Y)	---	d_0/d_4 BS²G	LinkerOH	K226	---	d_0 1516.759 d_4 1520.784	1516.758 1520.783	0.7 0.6	50x, 30min, T, [2)]
	220-230 (R)/SVCHLQKVFER/(Y)	---	d_0/d_8 DSA	LinkerOH	K226	---	d_0 1530.774 d_8 1538.822 d_0 1530.755 d_8 1538.818	1530.773 1538.823	0.6 0.6 11.7 3.2	100x, 30min, T, [2] 100x, 30 min, T, [1)]
	220-230 (R)/SVCHLQKVFER/(Y)	---	d_0/d_4 BS³	LinkerOH	S220, K226	---	d_0 1558.804 d_4 1562.830 d_0 1558.845 d_4 1562.878	1558.805 1562.830	0.6 0 25.6 30.7	100x, 30 min, T, [1)] 50x, 30 min, T, [1)]
	238-249 (Y)/DMLESIKKEVKG/(D) **238-245 x 246-249** (Y)/DMLESIKK/(E) x (K)/EVKG/(D)	---	sDST	LinkerOH or Intramol Interpept	K244, K245, K248 or K244 x K245, K244 x K248	---	1508.754 1508.756 1508.758	1508.751	1.9 3.3 4.6	20x, 15 min, AspN, [2)] 50x, 30min, AspN, [2)] 50x, 60 min, AspN, [2)]
	313-328 (K)/SLYNYIQQDTKGDYQK/(A)	---	d_0/d_4 BS²G	LinkerOH	K323, K328	---	d_0 2078.009 d_4 2082.020 d_0 2078.000 d_4 2082.024	2077.971 2081.996	18.2 11.5 13.9 13.4	50x, 30 min, T, [1)] 100x, 30 min, T, [1)]
	313-328 (K)/SLYNYIQQDTKGDYQK/(A)	---	d_0/d_8 DSA	LinkerOH	K323, K328	---	d_0 2091.957 d_8 2100.011	2091.987 2100.037	14.3 12.3	100x, 30 min, T, [1)]
	313-328 (K)/SLYNYIQQDTKGDYQK/(A)	---	d_0/d_4 BS³	LinkerOH	K323, K328	---	d_0 2120.028 d_4 2124.054 d_0 2120.018 d_4 2124.027 d_0 2120.045 d_4 2124.030	2120.018 2124.043	4.7 5.1 0 7.5 12.7 6.1	100x, 30min, T, [2] 100x, 30 min, T, [1)] 50x, 30 min, T, [1)]
~50kDa ANXA2 / p11	10-27 (K)/LSLEGDHSTPASAYGSVK/(A)	---	d_0/d_4 BS²G	LinkerOH	K27	---	d_0 1932.923 d_4 1936.941	1932.918 1936.943	2.5 1.0	50x, 30 min, T, [3), x)]
	10-27 (K)/LSLEGDHSTPASAYGSVK/(A)	---	d_0/d_8 DSA	LinkerOH	K27	---	d_0 1946.936 d_8 1954.986	1946.934 1954.984	1.0 1.0	100x, 30 min, T, [3), x)]
	37-48 (R)/DALNIETAIKTK/(G)	---	d_0/d_8 DSA	LinkerOH	**K46**	---	d_0 1444.793 d_8 1452.843	1444.789 1452.839	2.7 2.7	100x, 30 min, T, [3)]
	37-48 (R)/DALNIETAIKTK/(G)	---	d_0/d_4 BS²G	LinkerOH	**K46**	---	d_0 1430.777 d_4 1434.800	1430.774 1434.799	1.3 0.6	50x, 30 min, T, [3)]
	69-108 (Q)/DIAFAYQRRTKKELASALKSALSGHLETVILGLLKTPAQY/(D)	---	sDST	LinkerOH	K79, K80, K87, K103	---	4533.538 4533.529 4533.484 4533.491 4533.482 4533.516	4533.472	14.5 12.5 2.6 4.1 2.2 9.7	20x, 15 min, AspN, [2)] 20x, 30min, AspN, [2)] 20x, 60 min, AspN, [2)] 50x, 15 min, AspN, [2)] 50x, 30min, AspN, [2)] 50x, 60 min, AspN, [2)]

Group	Peptide / Sequence		Crosslinker	Type	Crosslinked Lys		Measured	Calc.	Δ ppm	Conditions
~ 50kDa ANXA2 / p11	69-108 (Q)/DIAFAYQRRTKKELASALKSALSGHLETVILGLLKTPAQY/(D)	---	sDST	LinkerOH 2x	K79, K80, K87, K103	---	4665.477	4665.484	1.5	20x, 60 min, AspN, 2)
	78-87 (R)/TKKELASALK/(S)	---	d_0/d_4 BS^{2}G	LinkerOH or Intramol interpept	K79, (K80)	---	d_0 1202.699 d_4 1206.724	1202.699 1206.724	0 0	50x, 30 min, T, 3)
	80-87 (K)/KELASALK/(S)	---	d_0/d_8 DSA	LinkerOH	K80	---	d_0 987.573 d_8 995.625	987.572 995.622	1.0 3.0	100x, 30 min, T, 3)
	109-124 (Y)/DASELKASMKGLGTDE/(D)	---	sDST	LinkerOH	K114, K118	---	1783.795	1783.790	2.8	50x, 15 min, AspN, 2)
	165-180 (G)/DFRKLMVALAKGRRAE/(D)	---	sDST	LinkerOH	K168, K175	---	1993.068 1993.067 1993.068	1993.065	1.5 1.0 1.5	50x, 15 min, AspN, 2) 50x, 30min, AspN, 2) 50x, 60 min, AspN, 2)
	196-204 (R)/DLYDAGVKR/(K)	---	d_0/d_4 BS^{2}G	LinkerOH	K203	---	d_0 1150.577 d_4 1154.597	1150.574 1154.599	2.6 1.7	50x, 30 min, T, 3)
	196-204 (R)/DLYDAGVKR/(K)	---	d_0/d_8 DSA	LinkerOH	K203	---	d_0 1164.590 d_8 1172.639	1164.590 1172.640	0 0.8	100x, 30 min, T, 3)
	196-204 (R)/DLYDAGVKR/(K)	---	d_0/d_4 BS^3	LinkerOH	K203	---	d_0 1192.619 d_4 1196.621 d_0 1192.578 d_4 1196.593	1192.621 1196.646	1.6 20.8 36.0 44.2	100x, 30min, T, 1) 50x, 30min, T, 1)
	196-205 (R)/DLYDAGVKRK/(G)	---	d_0/d_8 DSA	LinkerOH	K203, K205	---	d_0 1292.686 d_8 1300.737	1292.684 1300.734	1.5 2.3	100x, 30 min, T, 3)
	196-211 (R)/DLYDAGVKRKGTDVPK/(W)	---	d_0/d_8 DSA	LinkerOH or Intramol interpept	K203, K205	---	d_0 1889.997 d_8 1898.046	1889.997 1898.047	0 0.5	100x, 30 min, T, 3)
	196-211 (R)/DLYDAGVKRKGTDVPK/(W)	---	d_0/d_4 BS^3	LinkerOH or **Intramol interpept**	K203 x K205	---	d_0 1918.027 d_4 1922.054	1918.028 1922.053	0.5 0.5	100x, 30 min, T, 3)
	205-211 (R)/KGTDVPK/(W)	---	d_0/d_4 BS^3	LinkerOH	K205	---	d_0 900.503 d_4 904.529	900.504 904.529	1.1 0	100x, 30 min, T, 3)
	205-211 (R)/KGTDVPK/(W)	---	d_0/d_8 DSA	LinkerOH	K205	---	d_0 872.472 d_8 880.524	872.472 880.522	0 2.2	100x, 30 min, T, 3)
	220-230 (R)/SVCHLQKVFER/(Y)	---	d_0/d_4 BS^{2}G	LinkerOH	K226	---	d_0 1516.758 d_4 1520.784	1516.758 1520.783	0 0.6	50x, 30 min, T, 3)
	245-248 (K)/KEVK/(G)	---	d_0/d_8 DSA	LinkerOH	K245	---	d_0 631.366 d_8 639.416	631.366 639.416	0 0	100x, 30 min, T, 3)
	245-248 (K)/KEVK/(G)	---	d_0/d_4 BS^3	LinkerOH	K245, K248	---	d_0 659.398 d_4 663.424	659.397 663.422	1.5 3	100x, 30 min, T, 3)
	273-280 (R)/LYDSMKGK/(G)	---	d_0/d_4 BS^{2}G	LinkerOH	K278, K280	---	d_0 1071.505 d_4 1075.527	1071.503 1075.528	1.8 0.9	50x, 30 min, T, 3)
	279-283 (K)/GKGTR/(D)	---	d_0/d_4 BS^3	LinkerOH	K280	---	d_0 674.383 d_4 678.408	674.383 677.408	0 0	100x, 30 min, T, 3)
	284-289 (R)/DKVLIR/(I)	---	d_0/d_8 DSA	LinkerOH	K285	---	d_0 871.525 d_8 879.575	871.525 879.575	0 0	100x, 30 min, T, 3)
	284-289 (R)/DKVLIR/(I)	---	d_0/d_4 BS^{2}G	LinkerOH	K285	---	d_0 857.509 d_4 861.534	857.509 861.534	0 0	50x, 30 min, T, 3)
	295-303 (R)/SEVDMLKIR/(S)	---	d_0/d_8 DSA	LinkerOH	K301	---	d_0 1234.635 d_8 1242.689	1234.635 1242.685	0 3.2	100x, 30 min, T, 3)
	304-308 (R)/SEFKR/(K)	---	d_0/d_8 DSA	LinkerOH	K307	---	d_0 794.404 d_8 802.455	794.404 802.454	0 1.2	100x, 30 min, T, 3)
	304-308 (R)/SEFKR/(K)	---	d_0/d_4 BS^{2}G	LinkerOH	K307	---	d_0 780.388 d_4 784.413	780.389 784.414	1.2 1.2	50x, 30 min, T, 3)
	310-323 (K)/YGKSLYNYIQQDTK/(G)	---	d_0/d_8 DSA	LinkerOH	K312, K323	---	d_0 1848.904 d_8 1856.954	1848.901 1856.951	1.6 1.6	100x, 30 min, T, 3)
	313-328 (K)/SLYNYIQQDTKGDYQK/(A)	---	d_0/d_4 BS^{2}G	LinkerOH	K323	---	d_0 2077.976 d_8 2081.997	2077.971 2081.996	2.4 0.4	50x, 30 min, T, 3)
	313-328 (K)/SLYNYIQQDTKGDYQK/(A)	---	d_0/d_8 DSA	LinkerOH	K323	---	d_0 2091.992 d_8 2100.041	2091.987 2100.037	2.3 1.9	100x, 30 min, T, 3)
	313-328 (K)/SLYNYIQQDTKGDYQK/(A)	---	d_0/d_4 BS^{2}G	LinkerOH	K323, K328	---	d_0 2078.029 d_4 2082.009 d_0 2077.960 d_4 2082.011	2077.971 2081.996	27.9 6.2 5.2 7.2	50x, 30min, T, 1) 100x, 30min, T, 1)
	313-328 (K)/SLYNYIQQDTKGDYQK/(A)	---	d_0/d_4 BS^3	LinkerOH	K323, K328	---	d_0 2120.029 d_4 2124.013	2120.018 2124.043	5.1 14.1	100x, 30min, T, 1)
~ 75kDa (ANXA2)$_2$	196-204 (R)/DLYDAGVKR/(K)	---	d_0/d_4 BS^3	LinkerOH	K203	---	d_0 1192.622 d_4 1196.635	1192.621 1196.646	0.8 9.1	50x, 30min, T, 1)
	196-205 (R)/DLYDAGVKRK/(G)	---	d_0/d_4 BS^3	LinkerOH	K203, 205	---	d_0 1320.652 d_4 1324.638 d_0 1320.695 d_4 1324.679	1320.716 1324.741	48.4 77.7 15.9 46.8	20x, 30min, T, 1) 50x, 30min, T, 1)
	313-328 (K)/SLYNYIQQDTKGDYQK/(A)	---	d_0/d_4 BS^{2}G	LinkerOH	K323, K328	---	d_0 2077.995 d_4 2082.016 d_0 2077.954 d_4 2082.019	2077.971 2081.996	11.5 9.6 8.1 11.0	100x, 30min, T, 1) 50x, 30min, T, 1)
~85kDa (ANXA2)$_2$ / p11	69-108 (Q)/DIAFAYQRRTKKELASALKSALSGHLETVILGLLKTPAQY/(D)	---	sDST	LinkerOH	K79, K80, K87, K103	---	4533.488 4533.473 4533.461 4533.462	4533.472	3.5 0.2 2.4 2.2	20x, 15 min, AspN, 2) 20x, 30min, AspN, 2) 20x, 60 min, AspN, 2) 50x, 15 min, AspN, 2)
	109-124 (Y)/DASELKASMKGLGTDE/(D)	---	sDST	LinkerOH	K114, K118	---	1783.797 1783.795 1783.795	1783.790	3.9 2.8 2.8	50x, 15 min, AspN, 2) 50x, 30min, AspN, 2) 50x, 60 min, AspN, 2
	155-179 (L)/EKDIISDTSGDFRKLMVALAKGRRA/(E)	---	sDST	LinkerOH	K156, K168, K175	---	2529.550	2529.510	15.8	50x, 30min, AspN, 2)
100kDa (ANXA2 / p11)$_2$	37-48 (R)/DALNIETAIKTK/(G)	---	d_0/d_8 DSA	LinkerOH	K46, K48	---	d_0 1444.793 d_8 1452.845	1444.789 1452.839	2.7 4.1	100x, 30 min, T, 3)
	69-108 (Q)/DIAFAYQRRTKKELASALKSALSGHLETVILGLLKTPAQY/(D)	---	sDST	LinkerOH	K79, K80, K87, K103	---	4533.474 4533.489 4533.486	4533.472	0.4 3.7 3.0	20x, 15 min, AspN, 2) 20x, 30 min, AspN, 2) 20x, 60 min, AspN, 2)
	80-87 (K)/KELASALK/(S)	---	d_0/d_4 BS^3	LinkerOH	K80	---	d_0 1015.603 d_4 1019.630	1015.603 1019.628	0 1.9	50x, 30 min, T, 3)

100kDa (ANXA2 / p11)$_2$ →	165-180 (G)/DFRKLMVALAKGRRAE/(D)	---	sDST	LinkerOH 2x	**K168, K175**	---	2125.052	2125.071	8.9	20x, 15 min, AspN, [2]
	168-175 (R)/KLMVALAK/(G)	---	d_0/d_8 DSA	LinkerOH	**K168**	---	d_0 1001.609 d_8 1009.660	1001.606 1009.656	2.9 3.9	100x, 30 min, T, [3]
	196-204 (R)/DLYDAGVKR/(K)	---	d_0/d_8 DSA	LinkerOH	K203	---	d_0 1164.588 d_8 1172.641	1164.590 1172.640	1.7 0.8	50x, 30 min, T, [3]
	196-204 (R)/DLYDAGVKR/(K)	---	d_0/d_4 BS^3	LinkerOH	K203	---	d_0 1192.623 d_4 1196.647	1192.621 1196.646	1.6 0.8	50x, 30 min, T, [3]
	196-211 (R)/DLYDAGVKRKGTDVPK/(W) (R)/DLYDAGVKR/(K) x (R)/KGTDVPK/(W)	---	d_0/d_4 BS^3	LinkerOH or Intramol interpept	**K203 x K205**	---	d_0 1918.029 d_4 1922.050	1918.028 1922.053	0.5 1.5	50x, 30 min, T, [3]
	196-211 (R)/DLYDAGVKRKGTDVPK/(W)	---	d_0/d_8 DSA	LinkerOH or Intramol interpept	K203, K205, K211	---	d_0 1889.991 d_8 1899.042	1889.997 1899.047	3.1 2.6	100x, 30 min, T, [3]
	205-211 (R)/KGTDVPK/(W)	---	d_0/d_4 BS^3	LinkerOH	K205, K211	---	d_0 900.505 d_4 904.529	900.504 904.529	1.1 0	50x, 30 min, T, [3]
	220-230 (R)/SVCHLQKVFER/(Y)	---	d_0/d_8 DSA	LinkerOH	**K226**	---	d_0 1530.778 d_8 1538.827	1530.773 1538.823	3.2 2.5	100x, 30 min, T, [3]
	245-248 (K)/KEVK/(G)	---	d_0/d_8 DSA	LinkerOH	K245, K248	---	d_0 631.366 d_8 639.416	631.366 639.416	0 0	50x, 30 min, T, [3]
	279-283 (K)/GKGTR/(D)	---	d_0/d_4 BS^3	LinkerOH	K280	---	d_0 674.385 d_4 678.409	674.383 678.408	2.9 1.4	50x, 30 min, T, [3]
	284-289 (R)/DKVLIR/(I)	---	d_0/d_4 BS^3	LinkerOH	K285	---	d_0 899.556 d_4 903.581	899.556 903.581	0 0	50x, 30 min, T, [3]
	295-303 (R)/SEVDMLKIR/(S)	---	d_0/d_4 BS^3	LinkerOH	K301	---	d_0 1262.668 d_4 1266.693	1262.666 1266.691	1.5 1.5	50x, 30 min, T, [3]
	295-303 (R)/SEVDMLKIR/(S)	---	d_0/d_4 BS^3	LinkerOH	K301	---	d_0 1246.673 d_4 1250.700	1246.671 1250.696	1.6 3.1	50x, 30 min, T, [3]
	298-324 (V)/DMLKIRSEFKRKYGK SLYNYIQQDTKG/(D)	---	sDST	LinkerOH 2x	K301, K307, K309, K312, K323	---	3573.747	3573.753	1.6	20x, 15 min, AspN, [2]

Lysines, which are modified by a hydrolyzed cross-linker, do not provide direct information on interaction sites in protein complexes, but they yield valuable information about the solvent accessibilities of certain regions in proteins. The 14 lysines in ANXA2, which were unambiguously found to be modified by a hydrolyzed cross-linker, are displayed in **Figure 4.8**.
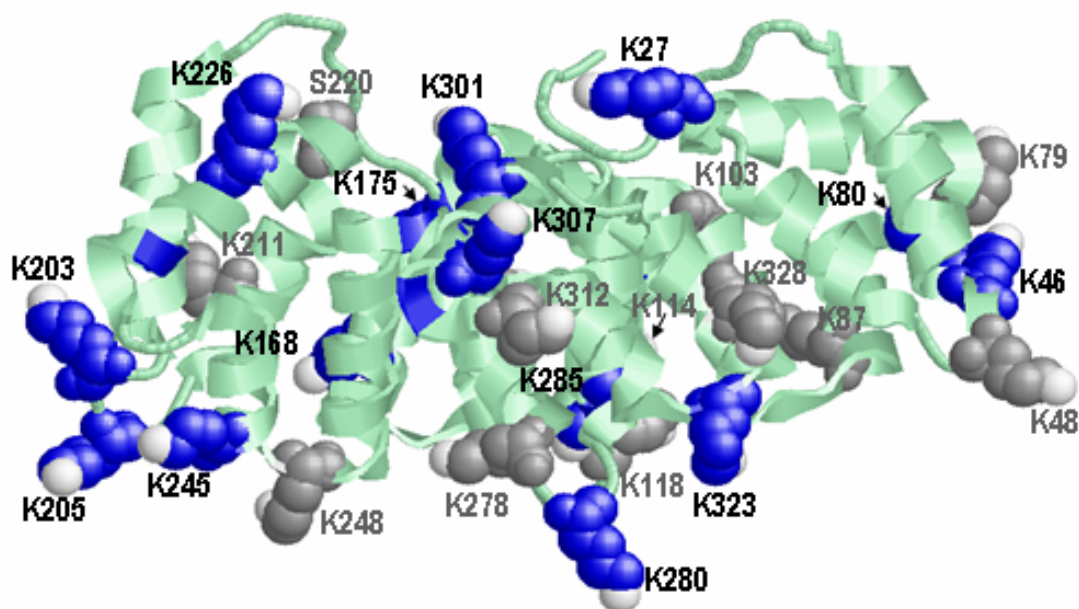
Figure 4.8: ANXA2 Residues Modified by a Partially Hydrolyzed Cross-Linker (PDB entry 1W7B) Visualized with RASMOL. Amino acids, which are modified by a hydrolyzed cross-linker, are displayed. Blue spheres: unambiguously identified lysines, gray spheres: ambiguously identified lysines and Ser-220. NZ and OG atoms are displayed in white.

As is visible from the structure of monomeric ANXA2 (PDB entry 1W7B), the side chains are located on the surface of ANXA2 and are therefore easily accessible for the cross-linking

reagent. This observation was validated using the GETAREA 1.1 software that calculates the solvent accessibility and declares amino acid residues that display more than 50% solvent accessibility to be exposed on the surface. Eleven out of the 14 unambiguously assigned lysines display solvent accessibilities between 52.2 and 100%, however, the remaining three lysines display solvent accessibilities below 50% (35.5 to 46.5 %).

For p11, only three out of eleven lysines were found to be potentially modified by a partially hydrolyzed cross-linker (**Figure 4.7**). The same amino acids were found to be involved in intramolecular cross-linking of p11, with calculated solvent accessibilities of 78.6, 92.4, and 60.4% for Lys-36, Lys-46, and Lys-53, respectively. Similarly, most amino acid residues involved in unambiguously assigned intramolecular cross-linked products of ANXA2 displayed calculated solvent accessibilities in the monomer of more than 50% (52.2 to 100%). For six lysine residues solvent accessibilities were below 50% (27.9 to 49.8%). Amino acid residues that were predicted to be buried within the molecule and, conclusively, are not accessible for the cross-linking reagent, were never (K-265, 22.7%), or less frequently (K-103, 27.9%), found to be modified. Interestingly, Ser-220, which exhibits only 10.1% solvent accessibility in monomeric ANXA2, was found to form an intrapeptide cross-link to Lys-226 (**Table 4-2**), and is potentially involved in an intermolecular cross-linked product (**Table 4-1**).

4.4 Low-Resolution Structure of A2t

In the first, low-resolution docking step, 20,000 models were generated employing the X-ray structures of ANXA2 (PDB entry 1W7B) and a p11 dimer plus two synthetic N-terminal ANXA2 peptides (PDB entry 1BT6) as starting structures. The models were filtered for the distance restraints listed in **Table 4-5**.

Table 4-5: Distance Constraints Used for Protein-Protein Docking

Three intermolecular cross-linked products and the gap between Ser-11 and Ala-20 in ANXA2[1)] were used as filters for screening the generated models.

	Distance Restraint	Filter [Å]	Cross-Linker (Å)	% of Models Eliminated
Low-Resolution Docking	Ser11$_{Anx}$ – Ala20$_{Anx}$	25	[1)]	86.3
CA-CA distances	Lys114 or 118$_{Anx}$ – Lys53 or 56$_{p11}$	25	sDST (6.4)	98.7
	Ser220 or Lys 226$_{Anx}$ – Lys93 or 95$_{p11}$	35	BS^3 (11.4)	99.2
	Lys168 or 175$_{Anx}$ – Lys53 or 56$_{p11}$	25	DSA (8.9)	99.4
High-Resolution Docking	Ser11$_{Anx}$ – Ala20$_{Anx}$	25	[1)]	
CA-CA distances	Lys118$_{Anx}$ – Lys56$_{p11}$	25	sDST (6.4)	
	Ser220 or Lys 226$_{Anx}$ – Lys93 or 95$_{p11}$	30	BS^3 (11.4)	
	Lys168 or 175$_{Anx}$ – Lys53 or 56$_{p11}$	25	DSA (8.9)	
NZ-NZ / NZ-OG distances	Lys118$_{Anx}$ – Lys56$_{p11}$	15	sDST (6.4)	
	Ser220 or Lys 226$_{Anx}$ – Lys93 or 95$_{p11}$	25	BS^3 (11.4)	
	Lys168 or 175$_{Anx}$ – Lys53 or 56$_{p11}$	15	DSA (8.9)	

The first restraint determines the CA atom of Ser-11 of the synthetic ANXA2 N-terminal peptide (11 amino acids, PDB entry 1BT6) to be at maximum 25 Å apart from the CA atom of the first residue (Ala-20) in the ANXA2 monomer (PDB entry 1W7B), in which the N-terminus is lacking. The missing eight amino acids span a distance of 12 Å in case they adopt an α-helical conformation, however, the value of 25 Å accounts for the possibility that this part of the protein is unstructured. The remaining three restraints comprise the identified intermolecular cross-linked residues (**Table 4-1**). CA-CA distances were chosen, considering the fact that for low-resolution docking only backbone atoms are used. The length and flexibility of the lysine side chains and the backbone flexibility were accounted for by adding ~ 15 Å to the spacer length of the respective cross-linker given by the manufacturer (www.piercenet.com). Since the four C-terminal amino acids of p11 are not included in the X-ray structure (1BT6) the restraint had to be determined for the CA atom of the ultimate amino acid of p11 (Lys-91), thus the value was increased to 35 Å. From the 20,000 computed models, only 0.6% passed all four filters and accordingly, 125 structures were used for subsequent high-resolution docking as detailed in the *Experimental Procedures*.

The A2t complex was generated by superimposition of a second ANXA2 molecule, taking advantage of the inherent symmetry of the p11 dimer. Of the eleven structures, three A2t structures displayed steric clashes and were eliminated. For the remaining eight structures the missing eight amino acids (aa 12-19 LEGDHSTP) of ANXA2 were modeled and amino acids were exchanged (P20A, P222C, D229E, R244K, Y316N) in order to account for the differences between human and porcine ANXA2. Ser-8 in the synthetic N-terminal ANXA2 peptide was exchanged to Cys-8, and the missing four C-terminal amino acids in p11 were added. The obtained tetramer models were compared to each other by overlay and grouped for similarity according to backbone root mean square deviation (rmsd) values below 10Å. **Figure 4.9 A** I-III shows the three model structures that represent the predominant conformations. They can be described as open, semi-open, and closed conformations (**Figure 4.9 A** I, II, and III, resp.). In each of the three models the two ANXA2 molecules reside on the same side of the p11 dimer, therefore strongly implying the existence of a heterooctameric state of the ANXA2 / p11 complex *in vivo*. The side chains of the amino acids that were identified to be involved in intermolecular cross-linked product formation are displayed.

Table 4-6 lists the CA-CA distances of the intermolecular cross-linked products between ANXA2 and p11 which were obtained from computational protein-protein docking. The first three in the list were those used as restraints for building the model of ANXA2 / p11.

Considering both the spacer lengths between 6.4 to 11.4 Å of the employed cross-linking reagents and the length and flexibility of lysine sidechains, the obtained CA-CA distances are well within the 25 Å cut-off value defined for the length of two lysine side chains plus one cross-linker molecule. This holds true for at least one potential cross-linked product within the cross-

linked ANXA2 and p11 peptides and is within this distance in at least one of the proposed models. However, for the last cross-link of the table, which was discussed as being ambiguously assigned in *chapter 4.3.1*, it is anticipated that it might be a false positive, as that the CA-CA distances obtained for all four possible cross-links in all three models, exceeds the distance of 25Å defined as maximum distance for docking.

Table 4-6: CA-CA Distances of Intermolecular Cross-Linked Products Obtained from Computational Docking of A2t.

Cross-Linked Product ANXA2 x p11	Distance [Å]		
	Model I CA	Model II CA	Model III CA
K118 x K56	22.2	22.3	23.8
S220 x K93	12.6	17.9	13.5
S220 x K95	9.6	10.8	13.1
K226 x K93	18.8	23.4	17.3
K226 x K95	16.4	18.2	14.9
K168 x K53	28.5	29.8	28.5
K168 x K56	28.4	27.6	29.3
K175 x K53	20.9	21.3	18.31
K175 x K56	21.8	20.1	19.7
K168 x K46	32.5	38.8	30.6
K175 x K46	23.7	28.7	20.8
K46 x K53	32.9	33.2	31.3
K46 x K56	37.0	41.5	35.7
K48 x K53	38.0	38.6	36.4
K48 x K56	42.0	36.5	30.3

In **Figure 4.9 B** I-III, the derived models for the heterooctameric complex are displayed, which might exert ANXA2 function *in vivo*. The dimensions (height x width x depth) for the three heterooctamer models are: 112 Å x 116 Å x 73 Å for model I, 118 Å x 101 Å x 87 Å for model II, and 122 Å x 103 Å x 71 Å for model III, respectively.

A

I
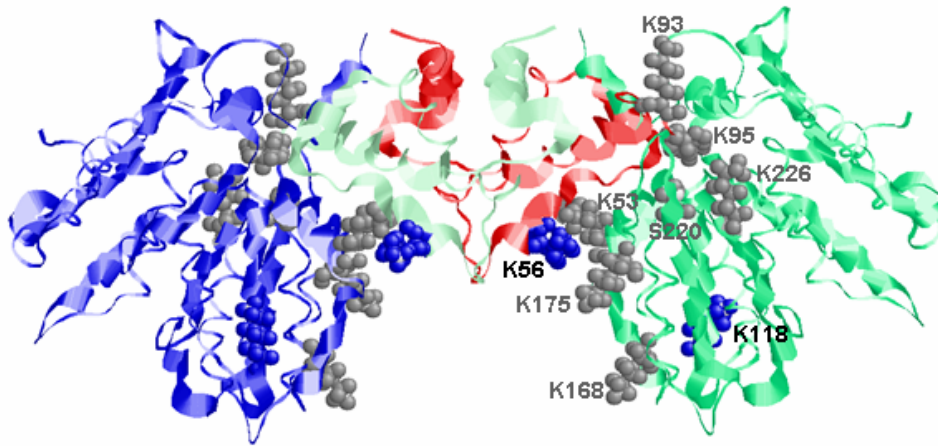
II
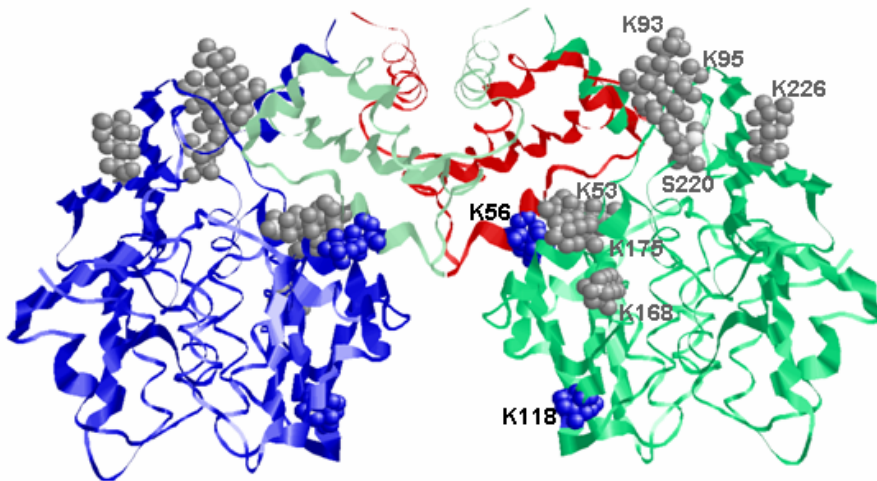
III
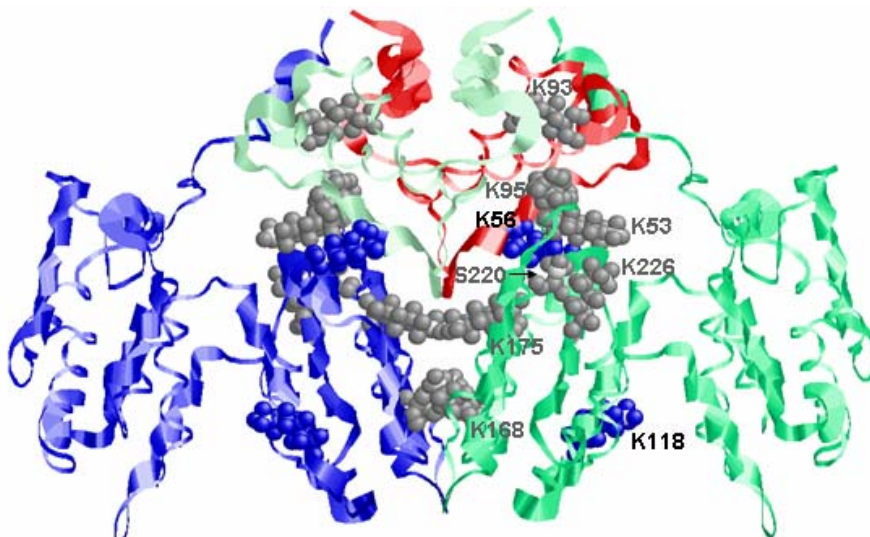
B

I
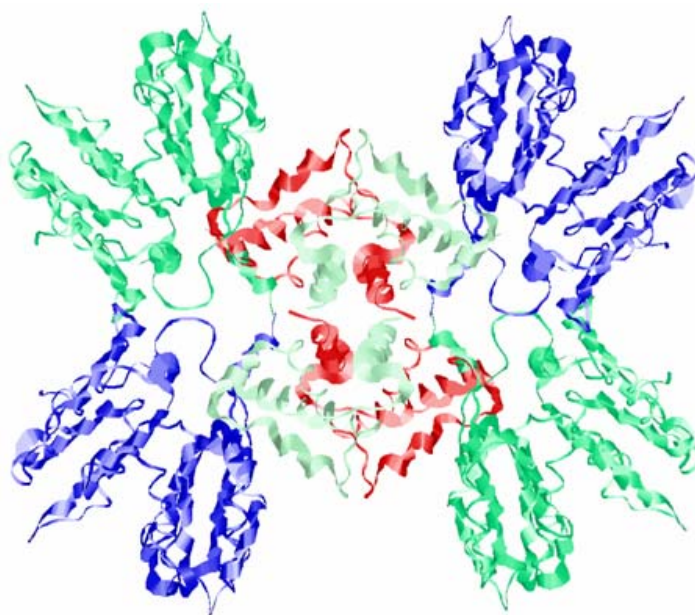
II
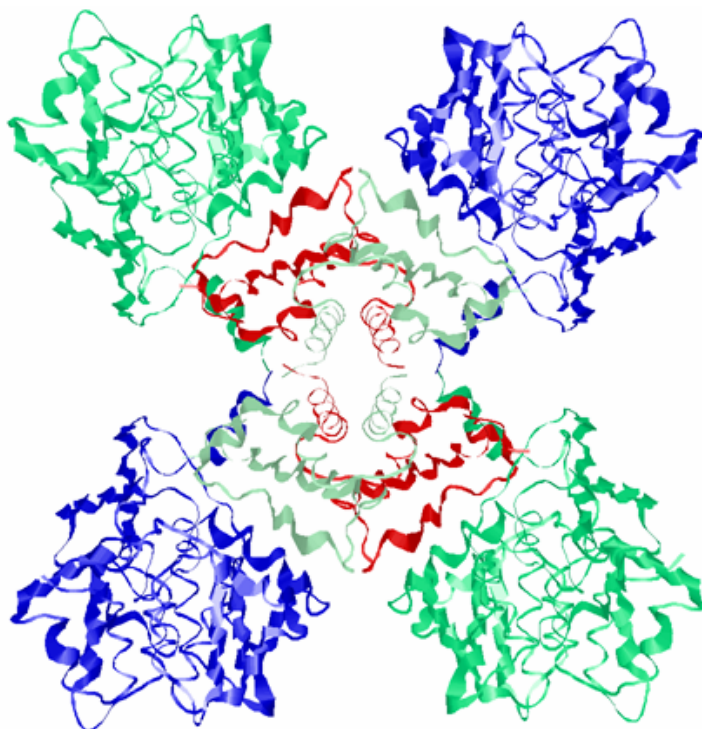
III
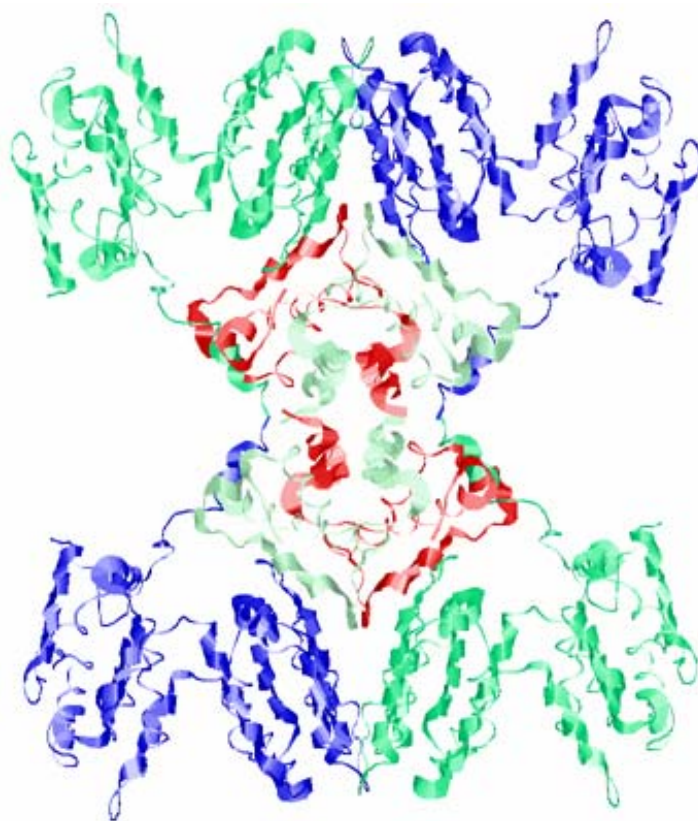
Figure 4.9: Models of ANXA2 / P11 Complexes. *A*, Models of the Heterotetrameric Complexes (A2t). I: open, II: semi-open, and III: closed conformation. The amino acids used as filters for generation of the models (**Table 4-5**) are displayed as blue (unambiguous) and gray (ambiguous) spheres. NZ and OG atoms are labeled white. B, Models of the Heterooctameric Complexes. I: open, II: semi-open, and III: closed conformation. Visualized with RASMOL.

4.5 Discussion

A structural model for the A2t complex was derived based on chemical cross-linking data in combination with high-resolution FTICR mass spectrometry and computational docking. Chemical cross-linking experiments of membrane-associated ANXA2 and A2t had been conducted before, using the amine-reactive cross-linkers dimethyl suberimidate (DMS) and dimethyl pimelimidate (DMP) [Faure *et al.*, 2002]. High-molecular weight aggregates as well as a variety of ANXA2 / p11 complexes with different stoichiometries had been found to be created during DMS cross-linking experiments in this previous study. The cross-linking experiments, which were conducted herein in the presence of 1 mM calcium, confirm and extend these findings. SDS-PAGE of the reaction mixtures from chemical cross-linking using four homobifunctional amine-reactive cross-linkers (**Figure 4.1 A** and **B**) revealed distinct bands for monomeric and dimeric ANXA2 and p11 in addition to cross-linked complexes between ANXA2 and p11 with varying stoichiometries, such as the ANXA2 / p11 heterodimer, the $(ANXA2)_2$ / p11 heterotrimer, the heterotetramer (A2t), and multimers. For clarification of the interaction sites between ANXA2 and p11, the bands corresponding to covalently attached subunits of A2t were cut out from the gel and subjected to enzymatic *in-gel* digestion before the peptide mixtures were analyzed by high-performance mass spectrometric techniques. Native gel electrophoresis under non-reducing conditions confirmed the existence of intact A2t or even a higher molecular weight aggregate (**Figure 3.9**). Thus, the various ANXA2 / p11 complexes that are visible in denaturing gel electrophoresis must originate from a high-molecular weight ANXA2 / p11 complex, in which the subunits fall apart, in case they are not connected by a covalent cross-link. Thus, each SDS-PAGE band corresponding to an ANXA2 / p11 complex with a specific stoichiometry will reveal information on the interaction sites within the native ANXA2 / p11 complex.

Conclusively, three intermolecular restraints between ANXA2 and p11 served as a basis for docking studies and for creating an A2t model in addition to the restraint given by the gap between the synthetic N-terminal ANXA2 peptide and the ANXA2 monomer lacking the N-terminus (**Table 4-5**). These cross-linking restraints were either confirmed by sequence data using MS/MS experiments and therefore are considered to be unambiguous or were obtained when employing isotope-labeled cross-linkers. The cross-linking restraint obtained with sDST, which was not confirmed by MS/MS data, was excluded for creating the A2t model. Eight unambiguously identified intramolecular restraints within ANXA2 and one ambiguous intramolecular cross-linked product within p11 gave hints on the conformation of both binding partners upon complex formation (**Table 4-2**). 14 unique unambiguously assigned lysine residues, which have been modified by a partially hydrolyzed cross-linker, provide valuable information on the surface topologies of the complex constituting components (**Figure 4.8**).

From the extent of modification of lysine residues, solvent accessibilities of certain regions in the complex were estimated.

Two different A2t models have so far been proposed. The first one is based on a cryo-electron microscopy study of junctions between phospholipid vesicles coated with the ANXA2 / p11 complex in the presence of calcium [Lambert *et al*., 1997]. This complex assumes the p11 dimer to be in the center of the tetramer, while one ANXA2 molecule resides on either side of the p11 dimer. In the second model, two molecules of ANXA2 are located on one side of the p11 dimer [Waisman, 1995]. In order for A2t to bridge two membranes, the latter model presumes that two A2t molecules form an octameric structure, which is connected by the p11 dimer. More recently presented scanning force microscopy data (topographic images) allowed determining the height of the protein layer [Menke *et al.*, 2004], which favors the second model. For models of the ANXA2 / p11 octamer (**Figure 4.9 B**) determined by chemical cross-linking and mass spectrometry, the heights of the complexes range between ca. 112 and 122 Å, which is larger than the value of ca. 90 Å that has been proposed by the cryo-electron microscopy study [Lambert, *et al.*, 1997]. Another electron microscopy study, however, estimated the diameter of A2t to be 107 +/- 17 Å [Nakata *et al.*, 1990], which is in good agreement with the dimensions of our proposed structural models.

It has been assumed that p11 forms a tetramer by creating a disulfide bond of Cys-61 with a neighbouring p11 dimer molecule [Réty *et al.*, 1999]. Nevertheless, oxidation of S100 proteins seems to be limited *in-vivo* under mostly denaturing conditions, and one cannot rule out the possibility that disulfide linked p11 dimers are artefacts due to storage conditions of the protein. In our peptide mass fingerprint experiments under non-denaturing conditions, we did not find any hint on a disulfide bond between two Cys-61 of p11 dimers. Previous alkylation studies of p11 residues Cys-61 and Cys-82 had revealed modification of both cysteines [Johnsson *et al.*, 1990]. If only Cys-61 was alkylated, binding activity for ANXA2 was retained. These findings contradict the formation of a p11 tetramer, in which Cys-61 is involved in disulfide bridge formation [Réty *et al.*, 1999].

There has been a previous report on creating models for A2t and its corresponding heterooctamer using docking algorithms [Sopkova-de Olivera Santos *et al.*, 2000]. The models proposed here for A2t and the heterooctameric complex differ from so far existing models [Waisman, 1995, Réty 1999, Sopkova-de Olivera Santos 2000]. Here (**Figure 4.9 A** and **B**), the two ANXA2 molecules reside on the same side of the p11 dimer. In one proposed structure (**Figure 4.9 A**, III and **Figure 4.9 B**, III), denoted closed structure, the two ANXA2 molecules come very close to each other, which might lead to the formation of cross-linked products between two ANXA2 molecules. However, without labeling the protein itself, e.g., by [15]N [Taverner *et al.*, 2002], it is impossible to tell whether a cross-linked product connects lysines within one single or between two ANXA2 molecules - in case the gel band of A2t is investigated.

The findings support the hypothesis that a heterooctamer is created *in-vivo* and shed new light into the interaction between ANXA2 and p11.

5 Identification of Binding Partners of the Annexin A2 / P11 Complex by Chemical Cross-Linking

In the quest of the –*omics* sciences for holistic schemes, the identification of binding partners of proteins is certainly gaining in importance. In this regard, chemical cross-linking represents a useful means through covalent attachment of interacting proteins, thus making transient interactions observable.

The goal was to establish an approach for the identification of A2t binding partners employing chemical cross-linking techniques. Briefly, A2t was first biotinylated and, in case a two-step cross-linking approach was chosen, reacted with a heterobifunctional amine- and photoreactive cross-linking reagent via its amine-reactive NHS ester group (**Figure 5.1**). The such labeled A2t was added to a crude mucosal preparation and allowed to interact with potential binding partners before the photoreaction or reaction with the homobifunctional amine-reactive cross-linker ethylene glycol *bis*(succinimidyl succinate) (EGS) was initiated. Afterwards, A2t was reconstituted similarly as described in *chapter 3.1.2*, involving a final affinity purification step via the biotin moiety.
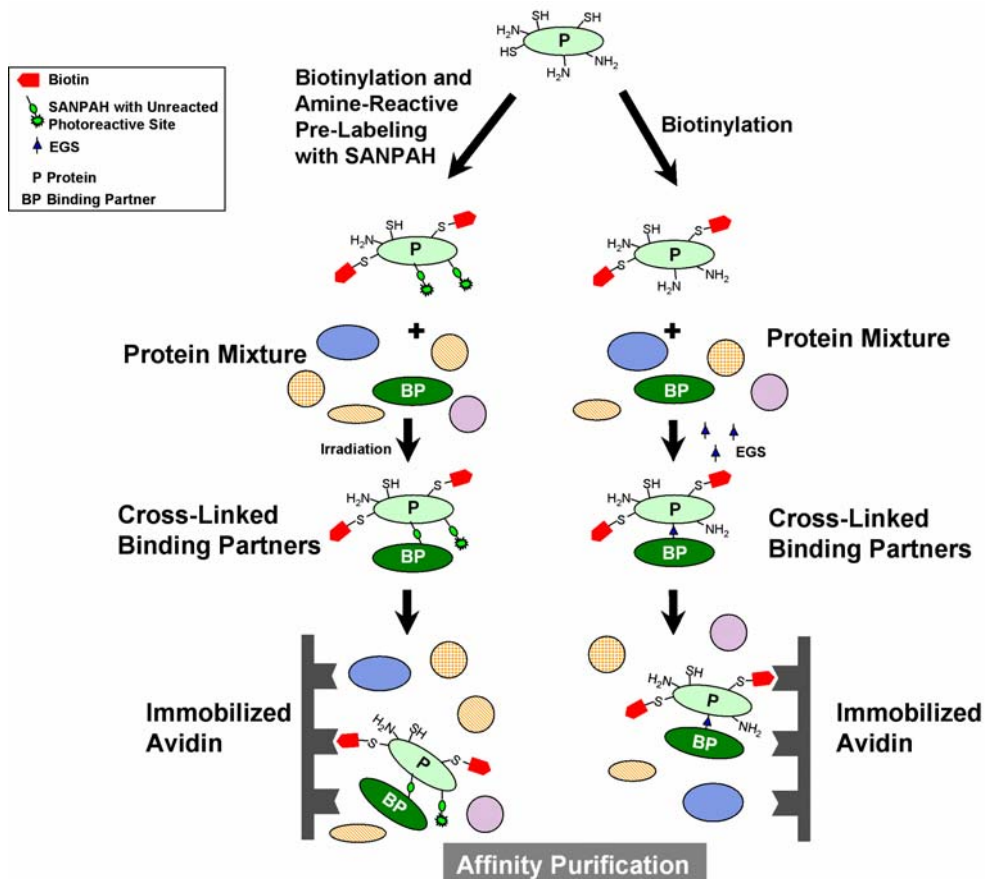
Figure 5.1: Analytical Strategy for the Determination of Protein Interaction Partners.

5.1 Biotinylation of A2t

For the purpose of affinity purification of cross-linked A2t with its binding partner complexes, purified A2t was modified with a biotin group. The sulfhydryl-reactive labeling reagent PEO-iodoacetyl biotin was reacted with A2t (1.5 µM) for 45 and 90 minutes at comparably high concentrations of 1, 2, and 4 mM, which were required due to the presence of 1 mM DTT, which was added during the purification process (chapter 3.1.2). The reaction was quenched by the addition of excess dithiothreitol (DTT).

Mass spectra that were very similar to that of non-labeled ANXA2 were obtained for ANXA2 that had been subjected to biotinylation (**Figure 5.2**). The acquisition of similar peptide patterns proves that addition of the labeling reagent does not induce undesirable by-products caused by side reactions, but yields well-defined, uniquely labeled products (**Figure 5.3**).
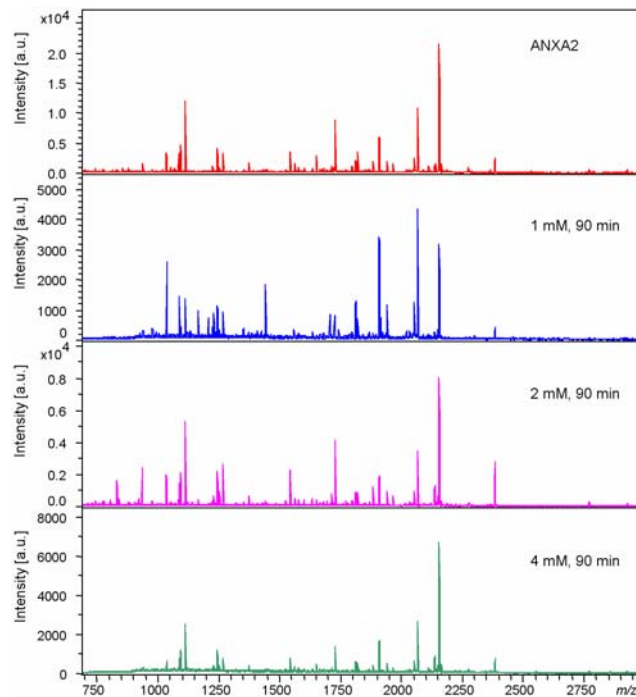
Figure 5.2: MALDI-TOF Mass Spectra of Biotinylated **ANXA2**. Tryptic digest of ANXA2 gel bands. Compared to the peptide pattern obtained for unmodified ANXA2 (top), the mass spectra of the labeled ANXA2 display similar peptide patterns.

Of the five cysteines present in ANXA2 (C8, C132, C222, C261, C334), C222 and C132 were frequently observed to be modified by the biotinylation reagent as determined by MALDI-TOFMS analyses (**Figure 5.3**) of tryptic *in-gel* digests of the ANXA2 monomeric gel band. **Table 5-1** lists the non-labeled and labeled peptides that were observed for the unmodified A2t and labeled A2t under different reaction conditions.

Table 5-1: Biotinylation of A2t. Theoretical *m/z* values for both non-labeled and labeled (asterisk) peptides are given. The reacted cysteines are underlined. X: observed, -: not observed, n.d.: not determined.

Sequence	[M+H]$^+$	A2t	1 mM, 45 min	2 mM, 45 min	4 mM, 45 min	1 mM, 90 min	2 mM, 90 min	4 mM, 90 min
220-226 SVCHLQK	814.424	X	n.d.	n.d.	n.d.	n.d.	X	n.d.
	1228.618 *	-	X	X	X	X	X	X
119-134 GLGTDEDSLIEIICSR	1720.854	X	-	-	-	-	-	-
	2135.036 *	-	-	-	X	X	X	X

Upon labeling of a cysteine residue with the employed biotinylation reagent, the respective peptide mass increases by 414.19 u. In **Figure 5.3 A** signals at 220-226 at *m/z* 1228.6 are displayed, corresponding to biotin-labeled ANXA2 peptide 200-226, which was obtained for all concentrations of reagent and reaction times employed (**Table 5-1**). In the mass spectrum of unmodified ANXA2 this signal was not observed, but instead the corresponding unmodified peptide was observed at *m/z* 814.4 (**Figure 5.3 B**). With the exception of MALDI-TOF analysis of A2t labeled with 2 mM reagent for 90 minutes, for which the signal at *m/z* 814 was not observed, the low-mass gate was set to *m/z* 900. Therefore, the signal at *m/z* 814 was not in the detection range. The second peptide that contained a labeled cysteine is that of ANXA2 amino acid sequence 119-134 with an *m/z* value of 2135.04 for the labeled species (**Figure 5.3 C**) and *m/z* 1720.9 for the unmodified peptide (**Figure 5.3 D**). The unmodified peptide was only observed in the mass spectrum of non-labeled ANXA2 (**Table 5-1**). The labeled peptide was observed in most spectra except for, of course, the unmodified ANXA2, but was also not observed at 45 minutes reaction time with concentrations of 1 and 2 mM PEO-iodoactyl biotin.

A

B
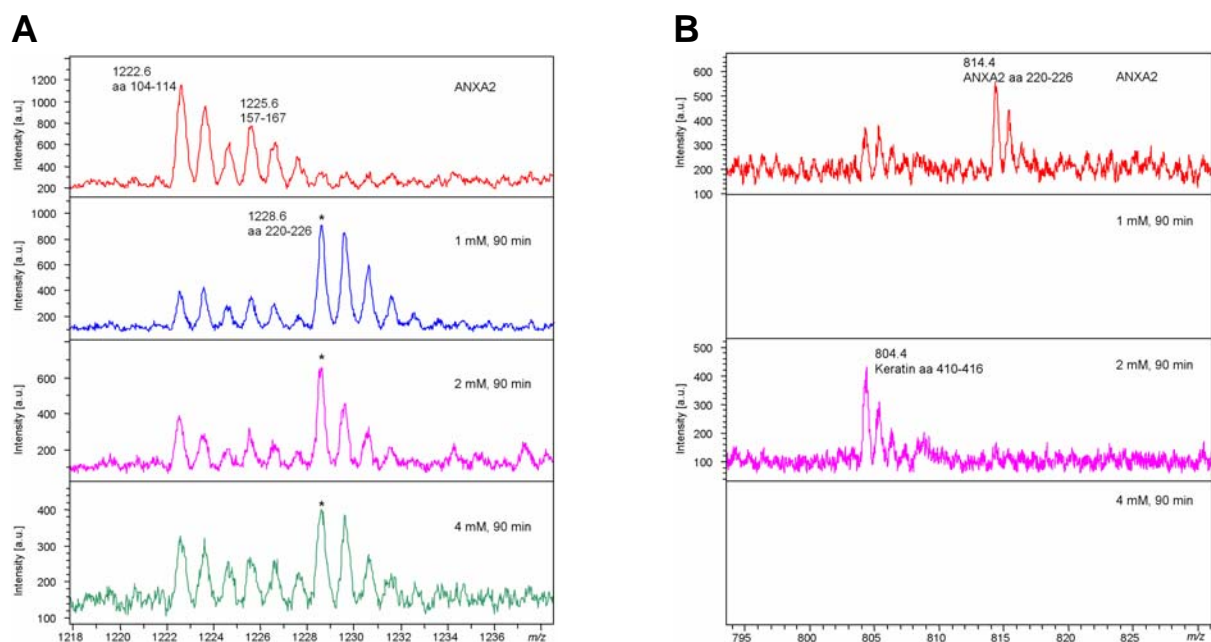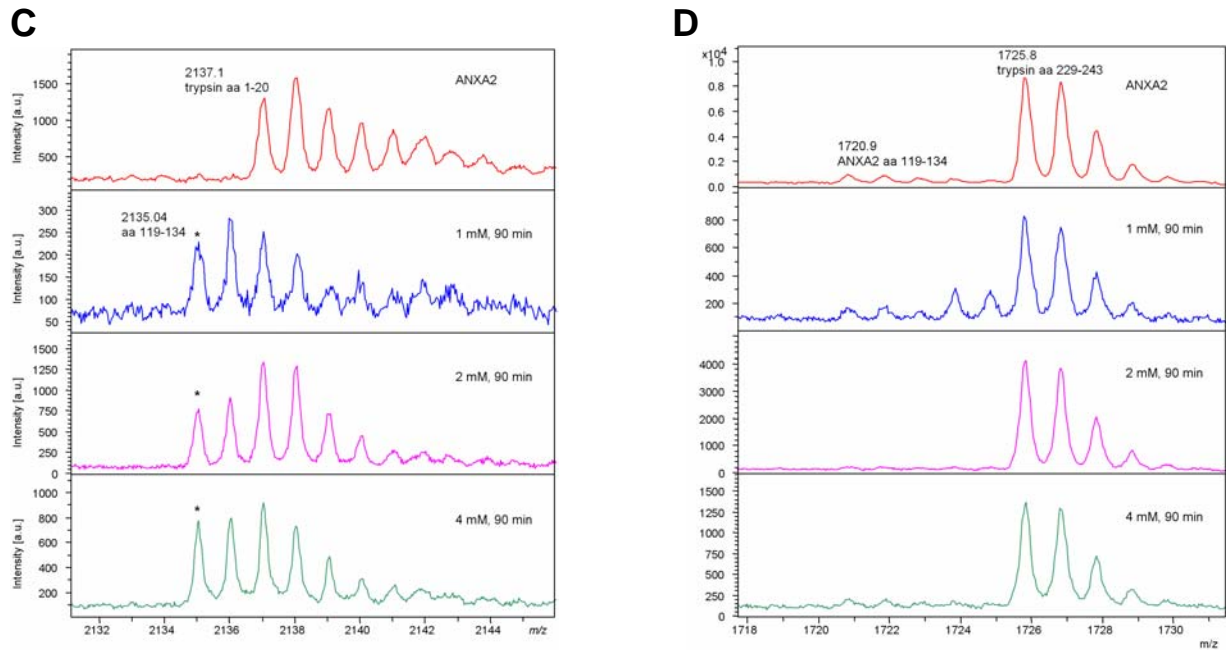
Figure 5.3: Signals of Biotinylated ANXA2 Peptides. Top to bottom: unmodified ANXA2; 1 mM reagent, 90 min reaction time; 2 mM reagent, 90 min reaction time; 4 mM reagent, 90 min reaction time. *A*, ANXA2 peptide 220-226 modified at Cys-222 (asterisk). Additionally, ANXA2 peptides at *m/z* 1222 and 1225 were observed in all samples. *B*, Unmodified ANXA2 peptide 220-226 at *m/z* 814. A The signal at *m/z* 804 represents a keratin peptide. *C*, ANXA2 peptide 119-134 modified at Cys-132 (asterisk). The signal at *m/z* 2137 represents an autolytic trypsin peptide. *D*, Unmodified ANXA2 peptide 119-134 at *m/z* 1720. The signal at *m/z* 1725 represents an autolytic trypsin peptide.

There might possibly be a third modified peptide (aa 1-9, N-terminally acetylated, *m/z* 1485.7), which was observed with very low signal intensities in two spectra (data not shown). However, the data are too ambiguous to adhere to this potential modification.

For subsequent binding partner search, a reaction time of 45 minutes in combination with 4 mM biotinylation reagent was chosen as this reaction condition yielded both labeled peptides described above within a short time (**Table 5-1**).

P11 was not probed for biotinylation as the tryptic peptides containing cysteines 61 and 82 are not easily amenable for mass spectrometric analysis. The generated peptides are either small (aa 57-62 at *m/z* 806) or long and hydrophobic (66-91). The latter tryptic peptide had never been observed in any mass spectrometric analysis conducted. Moreover, carbamidomethylation of cysteine residues of p11 had been observed to be incomplete. As it is the same type of chemical reaction for both carbamidomethylation and biotinylation it was concluded that biotinylation of p11 might be incomplete as well.

5.2 *In-situ* Cross-Linking of A2t / Binding Partner Assemblies

For *in-situ* cross-linking of A2t with its protein interaction partners, A2t was biotinylated with 4 mM biotin for 45 minutes. In case the heterobifunctional amine- and photoreactive cross-linker

SANPAH was used, the cross-linker was simultaneously added and also allowed to react for 45 minutes.

Prior to addition of pre-labeled A2t, the mucosal scrapings were thoroughly homogenized and washed in the presence of Ca^{2+} and DTT. Detergent was not employed to avoid loss of potential membrane-associated A2t interaction partners. Then, pre-labeled A2t was added and allowed to interact with potential interaction partners for 20 minutes. For covalent attachment of A2t with its binding partners, the sample containing the photoreactive cross-linking reagent *N*-succinimidyl-6-[4´-azido-2´-nitrophenylamino] hexanoate (SANPAH), was UV-irradiated, while EGS was added to sample preparation B (**Figure 5.1**). Extraction of cross-linked proteins was accomplished by thorough cell disruption, addition of EGTA and centrifugation. The supernatant containing complexes between biotin-labeled A2t and its binding partner was purified by affinity purification on avidin beads as detailed in the *Experimental Procedures*.

Figure 5.4 displays SDS-PAGE analysis of aliquots, which were collected during the just described cross-linking procedure.
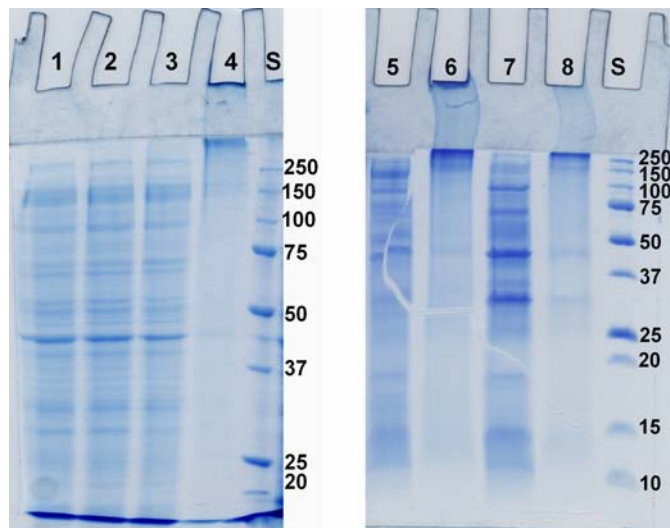
Figure 5.4: SDS-PAGE of A2t / Binding Partner *In-Situ* Cross-Linking. Lanes 1-9: *see text*; S: molecular weight standard (kDa).

Lanes 1 and 2 contain the supernatants of homogenized mucosa with biotin- and SANPAH-modified A2t (preparation A, lane 1) and homogenized tissue with biotin-labeled A2t (preparation B, lane 2), respectively. Lanes 3 and 4 represent preparations A and B after initiation of the cross-linking reaction, respectively. Following homogenization and centrifugation, the cytosolic fractions of preparations A and B (lanes 5 and 6, respectively) were obtained. The membrane fraction of preparations A and B (lanes 7 and 8, respectively) were obtained after adding the detergent SDS to the pellet and repeating the homogenizing and centrifugation step in order to obtain membrane-bound proteins in the supernatant.

Evidently, EGS induced excessive aggregation of proteins (lanes 4, 6, and 8), while no aggregation was visible for the photoreactive cross-linker SANPAH (lanes 3, 5, and 7).

5.3 Mass Spectrometric Identification of A2t Binding Proteins

The samples obtained from biotin / avidin affinity purification were separated by one-dimensional gel electrophoresis (**Figure 5.5 left hand side**) and gel bands were excised and tryptically digested. Following MALDI-TOF mass spectrometric analysis, the contained proteins were identified by peptide mass fingerprint (PMF) analysis using MS and MS/MS data.

Only the preparation involving the heterobifunctional amine- and photoreactive cross-linking reagent SANPAH was analyzed, because EGS (preparation B) caused excessive aggregation of virtually all proteins present in the mucosal preparation. Thus, hardly any low-molecular weight proteins were present in neither the cytosolic nor the membrane fraction as observed by SDS-PAGE (**Figure 5.4,** lanes 7 and 9).

Figure 5.5 shows the gel electrophoretically separated affinity-purified samples. On the right hand side, excised gel bands are indicated. Lanes 1 and 2 display the cytosolic and membrane fraction, respectively.
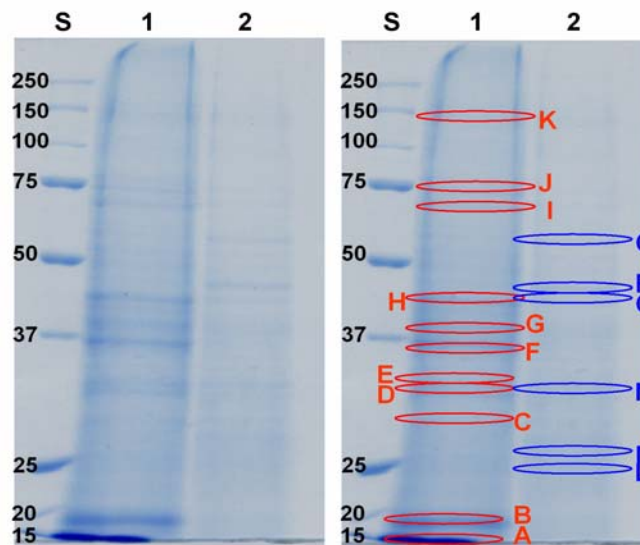
Figure 5.5: SDS-PAGE of Affinity-Purified A2t / Binding Partner Complexes Obtained from Photo-Cross-Linking. Cytosolic (lane 1) and membrane (lane 2) fractions of the preparation involving SANPAH as cross-linker. S: molecular weight standard [kDa]. The excised gel bands are indicated.

Table 5-2 lists all the proteins identified from the excised gel bands (**Figure 5.5**). Most notably, yet not surprisingly, the cytoskeletal components actin and cytokeratin are frequently identified. In general, analysis with the MALDI-TOF/TOF instrument (Ultraflex III) yielded higher MOWSE scores due to sequence data and superior mass accuracy compared to the MALDI-TOF mass

spectrometer (Autoflex I). The fact that ANXA2 was identified after the affinity purification step proves efficient and selective enrichment of biotinylated A2t by the employed protocol. Most of the identified proteins possess molecular weights close to the apparent molecular weights observed in SDS-PAGE or constitute fragments of the respective proteins.

Table 5-2: Identified A2t Binding Proteins.

	Gel Band	at kDa	Autoflex I			Ultraflex III		
			Identified Protein(s)	MW [kDa]	Score, no. of matched peptides	Identified Protein(s)	MW [kDa]	Score, no. of matched peptides, no. of ions for MS/MS
Cytosolic fraction	A	15	cytoskeletal beta-actin (fragment) pig	44.8	69, 9	cytokeratin-8 bovine galectin (L36) porcine	42.4 35.8	307, 24,5 120, 6, 2
	B	< 20	---	---	---	cytokeratin-19 bovine	43.9	349, 20, 5
	C	> 25	actin, cytoplasmic (beta-actin, fragment, pig)	44.7	86, 7	actin, cytoplasmic (beta-actin)	41.7	290, 14, 6
	D	> 25	ANXA4, pig	35.7	129, 12	ANXA4	35.7	354, 28, 7
	E	> 25	ANXA2, pig	38.5	87, 7	ANXA2 human	38.5	60, 7, 2
	F	37	---	---	---	cytokeratin-19	43.9	447, 28, 6
	G	> 37	MHC class I antigen, pig fragment leucocyte antigen, pig MHC class I antigen, pig	25 40.3 40. 5	108, 12 92, 13 85, 12	cytokeratin-20	48.5	170, 14, 3
	H	> 37	---	---	---	actin, cytoplasmic (beta-actin)	41.7	463, 22, 7
	I	< 75	Ig heavy chain, fragment cytokeratin-8, fragment	13.4 35.1	74 69	serum albumin pig	69.6	113, 23, 3
	J	75	GRP78 (BiP, Hsp70 family) fragment, pig GRP 78 lamin A/C	70.9 72.2 74	256, 20 237, 19 88, 19	GRP78 (glucose-regulated precursor, 78 kDa)	72.3	185, 13, 4
	K	150	---	---	---	---	---	---
Membrane fraction	L	25	---	---	---	---	---	---
	M	> 25	---	---	---	ANXA2 human	38.5	86, 6, 3
	N	> 25	ANXA2 pig	38.5	162, 21	ANXA2 human	38.5	112, 8, 2
	O	> 37	---	---	---	creatine kinase	46.9	165, 8, 4
	P	< 50	---	---	---	epoxide hydrolase (mEH), pig	52.4	115, 11, 2
	Q	> 50	protein disulfide isomerase A3 precursor	56.7	96, 13	protein disulfide isomerase A3 precursor	56.7	146, 15, 1

5.4 Discussion

The chemical cross-linking approach, which was employed for the *in-situ* determination of A2t binding partners, represents an overall promising strategy. However, further improvements are required to ensure validated results.

It was shown that initial biotinylation of the protein, for which interaction partners are to be determined, is by all means feasible and allows for a specific enrichment of later on cross-linked complexes by affinity purification using avidin beads. The optimum pH conditions for the reaction of PEO-iodoacetyl biotin is between 7.5 and 8.5. Side reactions with primary amino groups at pH > 10 and with histidine at pH < 7 are possible, but were not observed at pH 7.5. The two-step cross-linking strategy involving an initial amine-specific pre-labeling of A2t and a subsequent photoreactive cross-linking step has proven reasonably efficient for A2t binding partner complex formation. On the other hand, employing EGS as cross-linking reagent proved inapplicable, because of immediate aggregation upon addition of EGS to the mucosal preparation. As is visible from the gel (**Figure 5.4** lanes 5, 7, and 9), only high-molecular weight aggregates were obtained with EGS. No aggregation was observed when applying the sequential cross-linking strategy with the heterobifunctional amine- and photo-reactive cross-linker SANPAH (**Figure 5.4** lanes 4, 6, and 8). However, photocross-linking seemed to be hampered, because of the pulpy texture of the tissue preparation being impermeable for UV-irradiation, which possibly led to diminished photocross-linking efficiency.

The gel electrophoretically separated proteins obtained from two-step cross-linking and subsequent affinity purification were analyzed by MALDI-TOFMS and MALDI-TOF/TOFMS. In general, analyses with the MALDI-TOF/TOF instrument (Ultraflex III) yielded higher MOWSE scores due to MS/MS data and better mass accuracy. Most of the identified proteins possess molecular weights close to the apparent molecular weights observed in SDS-PAGE.

The majority of identified proteins constitute structural components of the cytoskeleton. A2t is well-known for its interaction with for example actin [Rescher & Gerke, 2004]. Two proteins (ANXA4 and galectin (L36)) identified earlier during the A2t purification procedure (**Figure 3.5**) were identified again. It is thus unclear whether these proteins constitute 'real' interaction partners or whether they had non-specifically bound to the avidin beads. On the other hand, galectin-4 (L36) and ANXA2 had been found, among other proteins, in lipid rafts of intestinal mucosal preparations, with these two proteins being present predominantly in the low-density regions of lipid rafts [Nguyen *et al.*, 2006].

Both galectins and heat shock proteins have been described as annexin A2 binding proteins. GRP78 (glucose-related protein) belongs to the family of heat shock proteins and is associated with the endoplasmic reticulum. Another protein of the Hsp family, Hsp90, has been found to be involved in the translocation of ANXA2 to the surface of endothelial cells [Lei *et al.*, 2004].

Lehner *et al*. [2003] reported on efficient membrane protein isolation strategies and identified among others protein disulfide isomerase A3 and ANXA2 in the membrane fraction and cytokeratin 8 in the soluble fraction. Furthermore, porcine serum album was identified in the membrane fraction hinting towards incomplete removal of cytosolic proteins from membrane-associated proteins. Conclusively, the proteins identified cannot unambiguously be considered 'real' binding partners of A2t.

The developed strategy, however, will serve as basis for a future development of protocols for the identification of so far unknown partners for a given protein.

6 Materials and Methods

6.1 Materials and Instrumentation

6.1.1 Proteins and Peptides

Angiotensin I	Sigma
Angiotensin II	Sigma
Bovine serum albumine (BSA)	Sigma
Peptide Calibration Standard	Bruker Daltonik
Calmodulin, high purity (*Bos taurus*, brain)	Calbiochem
Carbonic Anhydrase EC 4.2.1.1	Sigma
Chymotrypsin (EC 3.4.21.1, *Bos taurus*), sequencing grade	Roche Diagnostics
Cytochrome c (*Gallus gallus*)	Sigma
Endoproteinase AspN (EC 3.4.21.33, *Pseudomonas fragi* mutant), sequencing grade	Roche Diagnostics
Endoproteinase LysC (EC 3.4.21.50, *Lysobacter enzymogenes*), sequencing grade	Roche Diagnostics
Endoproteinase GluC (EC 3.4.21.19, *Staphylococcus aureus V8*) sequencing grade	Roche Diagnostics
Immunoglobulin A	Sigma
Lactate dehydrogenase	Sigma
Luteinizing hormone releasing hormone (LHRH)	Sigma
Lysozyme, chicken egg white	Sigma
Melittin, synthetic	Calbiochem
Monoclonal Anti-Annexin II, Clone CPI-50-5-1 (*Mus musculus*)	ICN Biomedicals
Myoglobin (*Equus caballus* ,heart)	Sigma
Phosphorylase b	Sigma
Somatostatin	Sigma
Substance P	Sigma
Trypsin (EC 3.4.21.4, *Bos taurus*), sequencing grade	Roche Diagnostics

6.1.2 Cross-Linking and Labeling Reagents

(+)-Biotinyl-iodoacetamidyl-3,6-dioxaoctanediamine (PEO-iodoacetyl biotin)	Pierce
Bis(sulfosuccinimidyl) glutarate (BS^{2}G), d_0 and d_4	Pierce
Bis(sulfosuccinimidyl) suberate (BS^3), d_0 and d_4	Pierce
Disuccinimidyl adipate (DSA), d_0 and d_8	K. Mechtler, IMP, Vienna
Disulfosuccinimidyl tartrate (sDST)	Pierce
Ethylene glycol *bis*(succiimidly succinate) (EGS)	Pierce
1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC)	Pierce
N-Hydroxysulfosuccinimide (Sulfo-NHS)	Pierce
N-succinimidyl-6-[4'-azido-2'-nitrophenylamino]hexanoate (SANPAH)	Pierce

6.1.3 Chemicals

Acetic acid, glacial	Merck/VWR
Acetone, Uvasol® for spectroscopy	Merck/VWR
Acetonitrile (ACN), Uvasol® for spectroscopy	Merck/VWR
Acrylamide/Bis solution 40% (37.5:1)	Bio-Rad
α-Cyano-4-hydroxy cinnamic acid	Bruker Daltonik, Sigma
Ammonium acetate	Sigma
Ammonium hydrogencarbonate (NH_4HCO_3)	Sigma
Ammonium persulfate (APS)	Bio-Rad, Sigma
6-aza-2-thiothymine (ATT)	Sigma
Calcium chloride	Sigma
Complete, EDTA-free protease inhibitors	Roche Diagnostics
Coomassie-Brilliant-Blue R250 and G250	Sigma
DEAE, pre-swollen microgranular DEAE cellulose (DE52)	Whatman
2',5'-Dihydroxyacetophenone	Sigma
2,5-Dihydroxybenzoic acid (DHB)	Sigma
3,5-Dimethoxy-4-hydroxy cinammic acid (sinapinic acid)	Sigma
Dimethyl sulfoxide (DMSO)	Sigma
Dithiothreitol (DTT)	Sigma
E-64 (*N*-[*N*-(L-3-carboxyoxirane-2-carbonyl)-L-leucyl]-agmatine)	Roche Diagnostics
Ethylene glycol tetraacetic acid (EGTA)	Sigma
Formic acid (FA)	Sigma, Merck/VWR
GelCode® Glycoprotein Staining Kit	Pierce
Glycine	Sigma
Hydrochloric acid (HCl)	Sigma
4-Hydroxyazobenzene-2-carboxylic acid (HABA)	Sigma
Imidazole	Sigma
Iodoacetamide	Sigma
N-(2-Hydroxyethyl)piperazin-N´-(2-ethansulfonsäure) (HEPES)	Sigma
Isopropanol, Uvasol®, for spectroscopy	Merck/VWR
2-Mercaptoethanol	Bio-Rad
Methanol, Uvasol®, for spectroscopy	Merck/VWR
2-(*N*-Morpholino)ethane sulfonic acid (MES)	Sigma
Native Sample Buffer, Laemmli	Bio-Rad
PeppermintStick™ phosphoprotein molecular weight standard	Invitrogen/Molecular Probes
Ponceau S	
Potassium chloride	Sigma
ProQ ® Diamond phosphoprotein gel stain	Invitrogen/Molecular Probes
Precision Plus Protein™ Unstained Standards (10-250 kDa)	Bio-Rad
Sample Buffer, Laemmli	Bio-Rad
sodium dodecyl sulfate (SDS) solution, 10%	Bio-Rad
Sodium chloride	Sigma, Roth
N,N,N´,N´-Tetramethylethylendiamine (TEMED)	Bio-Rad
Triluoro acetic acid (TFA)	Sigma
Tris / Glycine Running buffer (10x)	Roth
Trishydroxymethylaminomethane (Tris-Base)	Sigma
Trishydroxymethylaminomethane hydrochloride (Tris-HCl)	Sigma
Triton X-100 (octylphenolepoly(ethyleneglycolether)$_X$)	
	Pierce
UltraLink Immobilized Monomeric Avidin	

6.1.4 Instrumentation

MALDI-TOF Mass Spectrometers:
Voyager-DE™ RP Biospectrometry™ Workstation (Applied Biosystems)
Autoflex I (Bruker Daltonik)
Ultraflex III (Bruker Daltonik)

ESI-FTICRMS:
Apex II (Bruker Daltonics), 7T magnet, Nano-ESI source (Agilent)
LTQ-FT (ThermoScientific), 7T magnet, Nano-ESI source (Proxeon)

Nano-High-Performance Liquid Chromatography:
Ultimate™ Nano-HPLC System, (LC Packings / Dionex)
equipped with: Ultimate™ Micropump
Ultimate™ UV-Detector
Ultimate™ SWITCHOS II
Famos™ Micro-Autosampler

Ultimate™ 3000 Nano-HPLC System, (LC Packings / Dionex)

Fast Protein Liquid Chromatography (FPLC):
ÄKTA Explorer (GE Healthcare)

6.1.5 Miscellaneous Equipment and Consumables

Analytical balances	OHAUS Adventurer ARA520	OHAUS
	OHAUS Adventurer ARA640	
Bench-top shakers	Duomax 1030	Heidolph
	Titramax 101	
Bench-top centrifuge	MiniSpin	Eppendorf
Centrifuge	Avanti J-20 XP, rotor JLA 16.250	Beckman Coulter
	Optima LE-80K ultracentrifuge	
Desalting / Buffer Exchange	Microcon YM-3 and YM-10 centrifugal filter device (3 and 10 kDa cut-off)	Millipore
	ZipTips, C4 and C18	Millipore
	Centriprep YM-30, centrifugal filter device (30 kDa cut-off)	Millipore
	Dialysis membranes Spectra/Por®	Roth
Electrophoresis	Mini-PROTEAN 3 cell, POWERPAC 300 power supply, glass plates with integrated spacer, short plates, side-by-side casting stand, casting frames, electrode, clamping frame, sample loading guide, square-bottom plastic combs	Bio-Rad
Heating and drying oven	Heraeus T6 Function Line	Thermo Fisher
Imaging	Gel Image Scanner	Amersham
	PharosFX Molecular Imager System	Bio-Rad

pH-Meter	inoLab pH Level 1	WTW
Protein purification	Potter S homogenizer, homogenizer vessel, homogenizer cylinder, plunger	Sartorius
	Waring laboratory blender	Waring
Vacuum concentrator	Concentrator 5301 (SpeedVac)	Eppendorf
Water	DirectQ5^{TM} water purification system	Millipore
Western Blot equipment	Trans-Blot Semi-Dry cell	Bio-Rad

6.1.6 Software

ASAP	Automatic Spectrum Assignment Program, software for assigning MS peaklists generated from chemical cross-linking experiments, available at http://roswell.ca.sandia.gov/~mmyoung/asap.html
DataExplorer v. 4.0	Software for processing MALDI-TOF mass spectra (Applied Biosystems)
ExPASy Proteomics Server	Analytical tools for identification, sequence analysis, and tertiary structure prediction of proteins; database search; www.expasy.org
flexControl v. 2.2.19.0	Acquisition software for MALDI-TOFMS (Bruker Daltonik)
flexAnalysis v. 2.2	Software for processing of MALDI-TOF mass spectra (Bruker Daltonik)
GETAREA 1.1	Software that calculates the solvent accessible surface area of molecules; www.chem.ac.ru/Chemistry/Soft/GETAREA.en.html
GPMAW v. 7.01 and below	General Protein/Mass Analysis for Windows (Lighthouse Data, www.welcome.to/gpmaw); evaluation of peaklists from MS analysis for identification of cross-linked products and peptides
IsoFind	*In-house* developed tool for searching for distinct distances between two signals from a given MS peaklist and for calculating signal ratios based on signal intensity or peak area
Mammoth	Matching Molecular Models Obtained from Theory, program for sequence independent structure alignment of proteins
Mascot	Performs peptide mass fingerprint analysis from a given MS peaklist, www.matrixscience.com
Mascot Distiller v. 1.1, 2.0	Software for processing mass spectra (Matrix Science)
MS2 Assign	Tool for assigning peaklists from MS/MS experiments to a theoretical fragment library for cross-linked products, modified peptides, and unmodified peptides available at http://roswell.ca.sandia.gov/~mmyoung/ms2assign.html
Profound	Performs peptide mass fingerprint analysis for a given MS peaklist
Rasmol v. 2.7.3	Visualization and analysis of protein structures, www.openrasmol.org
Rosetta	Program for modeling, docking, etc of proteins; www. rosettacommons.org

UMAX scanner software	Operates UMAX scanner
Unicorn v. 4.10	Software for operating ÄKTA FPLC systems (Amersham Biosciences)
Voyager v. 5.1	Acquisition software for MALDI-TOFMS (Applied Biosystems)
VMD-Explorer v. 1.8.1	Visualization and analysis of protein structures (Theoretical and Computational Biophysics Group, NIH, www.ks.uiuc.edu)
XMASS vs, 5.0.10, 6.0 and 7.02	Software for acquisition and processing of ESI-FTICR mass spectra (Bruker Daltonics)
Xplor-NIH	Structure determination program, http://nmr.cit.nih.gov/xplor-nih/

6.1.7 List of Manufacturers

Agilent Technologies
Waldbronn
Germany

Applied Biosystems
850 Lincoln Drive
Foster City, CA 94404, USA

Bio-Rad
1000 Alfred Nobel Drive
Hercules, CA 94547, USA

Bruker Daltonics
Billerica, MA
USA

Bruker Daltonik GmbH
Bremen
Germany

Calbiochem
Schwalbach am Taunus
Germany

Eppendorf GmbH
Peter-Henlein-Str. 2
50389 Wesseling-Berzendorf
Germany

GE Healtcare
München
Germany

Heidolph Instruments
GmbH & Co. KG
Walpersdorfer Str. 12
91126 Schwabach, Germany

ICN Biomedicals, GmbH
Mühlgrabenstr. 12
53340 Meckenheim

Invitrogen / Molecular Probes
3 Fountain Drive
Inchinnan Buisness Park
Paisley PA4 9RF, UK

LC Packings / Dionex
Amsterdam
The Netherlands

Lighthouse Data
Odense
Denmark

Merck / VWR
Darmstadt
Germany

Millipore
Eschborn
Germany

New Objective
Woburn, MA
USA

Pierce Inc.
Rockford, IL
USA

Proxeon Biosystems
Odense
Denmark

Roche Diagnostics GmbH
Sandhofer Str. 116
68305 Mannheim, Germany

Roth (Carl-Roth GmbH & Co. KG)
Schoemperlenstr. 3-5
76185 Karlsruhe, Germany

Sartorius AG
Weender Landstr. 94-108
37075 Göttingen, Germany

Sigma-Aldrich Chemie GmbH
Eschenstr. 5
82024 Taufkirchen, Germany

Thermo Electron
Bremen
Germany

Waring Laboratory and Science
waringproducts.com

Whatman Inc.
27 Great West Road
Brentford, Middlesex
TW8 9BW, United Kingdom

WTW GmbH
Dr.-Karl-Slevogt-Strasse 1
D-82362 Weilheim

6.2 Experimental Procedures

6.2.1   Isolation and Purification of A2t

The ANXA2 / p11 heterotetramer (A2t) was purified from mucosa of pig (*Sus scrofa*) small intestines following a slightly modified version of the protocol of Gerke and Weber (1984).

Six fresh small intestines were obtained from the slaughterhouse Altenburg and processed immediately after evisceration. The small intestines were washed with 30 liters ice-cold imidazole buffer (10mM imidazole, 150 mM NaCl, pH 7.4) to remove the gut contents and were cut into pieces of about 50 cm length and were slit lengthwise. The thin mucosal layer was scraped off and was immediately frozen in liquid nitrogen for storage. Henceforth, all steps of the purification procedure were carried out at 4°C. Back in the laboratory, the frozen material was thawed, 1.5 l HEPES/Tris buffer (30 mM HEPES, 600 mM NaCl, 0.5% Triton X-100, 1 mM CaCl$_2$, pH 7.4, Tris was used for adjusting pH) was added, and the mucosa was mechanically disrupted using a Waring blender. In addition to protease inhibitors, the buffer contained 1 mM Ca^{2+}, which causes A2t binding to the membrane. The calcium concentration was raised to 2 mM and the homogenate was stirred for 15 minutes. After a centrifugation step (34,000 x g, 60 minutes, Avanti J-20 XP centrifuge, JLA 16.250 rotor, 4°C) the supernatant was discarded and the pellet containing A2t was retained. In total, the sample was washed and centrifuged three times with the above mentioned buffer containing the detergent Triton X-100, and was washed another three times without detergent. In a next step, the pellet was resuspended in 300 ml HEPES/Tris buffer (30 mM HEPES, 600 mM KCl) containing 10 mM ethylene glycol tetraacetic acid (EGTA) and the cells were thoroughly disrupted in a Potter homogenizer (Sartorius). Complexation of Ca^{2+} by EGTA released A2t from the membrane so that after ultracentrifugation (45,000 rpm, 60 minutes, Optima LE-80K ultracentrifuge, 4°C) the A2t-containing supernatant was retained. For preparation for DEAE (diethylaminoethyl cellulose, Whatman) anionic exchange chromatography, the supernatant was dialyzed over-night against two-times 15 l 20 mM imidazole / 10 mM NaCl / 0.5 mM EGTA buffer (pH 7.5). The dialyzed protein solution was applied onto the equilibrated DEAE column (column volume approx. 70 ml) and A2t was obtained in the flow-through. The obtained A2t was used without further purification. Chromatographic separation steps were conducted with the ÄKTA Explorer fast protein liquid chromatography (FPLC) system (GE Healthcare). The flow rate was manually adjusted and a one-step gradient of 10 mM and 1M NaCl was applied. UV absorption at 280 nm and conductivity were monitored.

Protease inhibitors *Complete EDTA-free* and E-64 as well as dithiothreitol (DTT, 1mM) were applied during the whole course of the purification. After each purification step aliquots were taken for monitoring the purification process and were stored at –20°C before SDS-PAGE and MS analyses were performed.

The purified A2t was concentrated using Centripreps YM-30 (Millipore, 30 kDa cut-off) and the buffer was at the same time exchanged against 20 mM HEPES / 150 mM NaCl / 0.5 mM EGTA, 0.5 mM DTT (pH 7.4). The purified A2t was then lyophilized for storage at –20°C.

6.2.2   Characterization of A2t

Purification of A2t was confirmed by Western blot analysis and A2t was characterized with respect to amino acid sequence and possibly existing posttranslational modifications.

6.2.2.1   Western Blot Analysis

Unstained gels from SDS-PAGE (chapter 6.2.6.2) were equilibrated in blot buffer (48 mM Tris, 39 mM glycine, 1.3 mM SDS, 20 % methanol (v/v), pH 9.2) for 20 minutes. Nitrocellulose membranes were cut to gel size and placed for three minutes in blot buffer. Filter paper was as well soaked in blot buffer. A sandwich composed of wet filter paper, nitrocellulose membrane, equilibrated gel, and another wet filter paper was placed on the bottom platinum anode of the TransBlot® Semi-Dry Transfer Cell (Bio-Rad) and the stainless steel cathode and safety cover were placed on top. Transfer of the proteins from the gel onto the nitrocellulose membrane was conducted at 10V (PowerPack300, Bio-Rad) for 30 minutes. Afterwards, the membrane was incubated in fixing solution (40 % methanol (v/v), 10 % trichloro acetic acid (TCA) (v/v)) for 20 minutes. Ponceau S staining was employed for confirming the protein transfer to the membrane. For blocking the membrane a 1% (w/v) bovine serum albumin solution in TTBS (Tween 20 (0.05% (v/v)) in TBS (20 mM Tris, 150 mM NaCl, pH 7.5)) was used, in which the membrane was incubated for 90 minutes under gentle agitation. The BSA / TTBS solution was discarded. Monoclonal anti-annexin II (clone CPI-50-5-1, 0.25 mg/ml, from *Mus musculus*, ICN Biomedicals) was diluted 1:5000 in TTBS and used for incubating the membrane (60 minutes). Unbound primary antibody was removed by three consecutive two-minute washing steps with TTBS and replaced by the secondary antibody anti-mouse IgG (diluted 1:10000 in TTBS, from goat (*Capra hircus*), Sigma product no. A3562)). This mouse-specific antibody is linked to the reporter enzyme alkaline phosphatase (AP). Incubation time was 45 minutes after which excess anti-body was again removed by two washing steps with TTBS and one washing step with 1x Development Buffer (Bio-Rad, AP Color Development kit). For colorimetric detection, 300 µl each of AP Color Reagents A (nitroblue tetrazolium, NBT) and B (5-bromo-4-cloro-3-indolyl phosphate, BCIP) (Bio-Rad) were mixed with 30 ml 1x Development Buffer. The nitrocellulose membrane was immersed in the color development solution for three minutes under gentle agitation. The reaction of BCIP with alkaline phosphatase results in a blue precipitate that is converted by NBT to result in purple color. Then the membrane was thoroughly rinsed with water and the air-dried membrane was imaged for documentation.

6.2.2.2 Amino Acid Sequences

The amino acid sequence of porcine ANXA2 was determined by ESI-FTICRMS (chapter 6.2.8.2) and MALDI-TOFMS (Autoflex I, chapter 6.2.7.3) of enzymatic digests (chapter 6.2.6.1). The obtained MS peaklists were compared to Swiss-Prot database (www.expasy.org) entries available for *Homo sapiens* (P07355), *Bos taurus* (P04272), *Canis familiaris* (Q6TEQ7)*, Sus scrofa* (P19620, fragment of amino acids 1-91), and porcine amino acid sequences suggested by V. Gerke (unpublished) and M. (Ph.D. thesis, 2005). ANXA2 amino acid sequence assessment was performed using the Expasy *Find Pept* tool (chapter 6.2.11.3). Experimentally obtained p11 peptide masses were compared to Swiss-Prot entry P04163.

6.2.2.3 Molecular Weight Determination of ANXA2 and P11

Linear MALDI-TOFMS analysis (Autoflex I) in the positive ionization mode of the purified A2t complex was performed for determination of the molecular weights of ANXA2 and p11 (chapter 6.2.7.3). Several MALDI matrices (sinapinic acid, DHB, ATT, HABA, and acetophenone) were tested for obtaining optimum mass spectra of ANXA2.

6.2.2.4 Posttranslational Modifications

Phosphorylation Analysis

ANXA2 and p11 were tested for phosphorylation using the Pro-Q® Diamond Phosphoprotein Gel Stain Kit (Invitrogen/Molecular Probes), an *in-gel* fluorescent detection assay for phosphorylation of Tyr, Ser, and Thr. PeppermintStick™ phosphoprotein molecular weight (MW) standard contains a mixture of two phosphorylated (ovalbumin (45 kDa) and α-casein (23.6 kDa)) and four non-phosphorylated proteins (β-galactosidase (116.25 kDa), BSA (66.6 kDa), avidin (18 kDa), and lysozyme (14.4 kDa)), thus functioning both as molecular weight standard and as positive and negative controls. Following separation of A2t (20 µl of a 5.5 µM solution) and the MW standard by gel electrophoresis, without subsequent Coomassie Brilliant Blue-staining, the gel (5% stacking / 12% resolving) was incubated twice in fixing solution (50% (v/v) methanol / 10% (v/v) acetic acid) for 30 minutes. The fixing solution was removed and the gel was immersed in H_2O for 10 minutes. Washing with H_2O was repeated twice for thorough removal of residual methanol and acetic acid. The gel was then incubated in ~ 60 ml of the staining solution for 90 minutes in the dark. Afterwards, the gel was destained in 100 ml destaining solution for 30 minutes, again protected from light. Destaining was repeated two more times with fresh solution and afterwards the gel was thoroughly immersed in H_2O. Visualization of stained protein bands was achieved with the PharosFX Molecular Imager System (Bio-Rad). The excitation and emission maxima of the Pro-Q® Diamond stain are ~555

nm and ~580 nm, respectively. Following documentation, the gel was stained with Coomassie Brilliant Blue.

Glycosylation Analysis

ANXA2 and p11 were tested for glycosylation using the GelCode® Glycoprotein Staining Kit (Pierce). Horseradish peroxidase and soybean trypsin inhibitor served as positive and negative controls, respectively. Gels from electrophoretic separation of A2t, positive and negative controls, and MW standard were fixed by 50% (v/v) methanol for 30 minutes. The solution was replaced by 100 ml 3% (v/v) acetic acid and incubated for ten minutes. This washing step was repeated once more, before the gel was transferred to 25 ml of Oxidizing Solution (periodic acid) and was incubated for 15 minutes, oxidizing glycols to aldehydes. The gel was washed three times for five minutes with 100 ml of 3% (v/v) acetic acid and was then immersed in 25 ml of GelCode® Glycoprotein Stain for 15 minutes. Afterwards the gel was transferred into 25 ml Reducing Solution and incubated for five minutes, after which the gel was extensively washed with 3% (v/v) acetic acid and H_2O. Glycoproteins appeared as magenta bands on the gel. The GelCode® Glycoprotein Staining Kit reagents were prepared as specified by the manufacturer.

6.2.3 Chemical Cross-Linking

6.2.3.1 *Cross-Linking Reactions of the Calmodulin / Melittin Complex*

For chemical cross-linking of the calmodulin / melittin complex the zero-length cross-linking reagent 1-ethyl-3(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) in combination with sulfo-*N*-hydroxysuccinimide (sNHS) was employed, as well the homobifunctional amine-reactive reagents sulfo-disuccinimidyl tartrate (sDST) and *bis*-sulfosuccinimidyl suberate (BS[3]).

EDC / Sulfo-NHS

An equimolar mixture of CaM and melittin (10µM each, final concentration) containing 1mM $CaCl_2$ (in 100mM MES buffer, pH 6.5) was incubated at room temperature for 20 min on a shaker. The zero-length cross-linker EDC was dissolved in water shortly prior to addition and added in 500-, 1000-, and 2000-fold molar excess over the protein/peptide mixture. Simultaneously, sNHS, dissolved in water, was added in 500-fold molar excess to all three EDC-containing reaction mixtures, thus resulting in EDC/sNHS ratios of 1:1, 2:1, and 4:1, respectively. As a negative control, the protein mixture was incubated without cross-linking reagent. The final volume was 1 ml. For quenching the reactions, 200µl-aliquots were taken from the reaction mixtures after 5, 15, 30, 60, and 120 min, and DTT (40mM final concentration) was added and the samples were stored at –20°C.

Table 6-1: Composition of the CaM / Mel Reaction Mixture with the Cross-Linker EDC / sNHS

	I	II	III	IV	Final concentration
EDC/Sulfo-NHS ratio	-	1:1	2:1	4:1	
0.4M EDC*	0µl	12.5µl	25µl	50µl	I: 0; II: 5mM; III: 10mM; IV: 20mM
0.4M sNHS*	0µl	12.5µl	12.5µl	12.5µl	I: 0; II, III, IV: 5mM
1mg/ml Melittin**	28.5µl	28.5µl	28.5µl	28.5µl	10µM
1mg/ml Calmodulin**	167µl	167µl	167µl	167µl	10µM
0.1M MES (pH 6.5)	794.5µl	769.5µl	757µl	732µl	≈ 0.1M
100mM CaCl$_2$**	10 µl	10 µl	10 µl	10 µl	1mM Ca^{2+}

* in H$_2$O, ** in MES buffer

sDST and BS^3

For cross-linking experiments with the homobifunctional cross-linking reagents BS^3 or sDST, an equimolar mixture of CaM and melittin (10µM each, final concentration) containing 1mM CaCl$_2$ (20mM HEPES buffer, pH 7.4) was incubated at ambient temperature for 20 min. BS^3 or sDST was dissolved in dimethyl sulfoxide (DMSO) to avoid hydrolysis (stock solutions: 10mM, 50mM, and 100mM, respectively) and 10µl were added to result in a 10-, 50-, and 100-fold molar excess over the proteins. The total volume of the reaction mixture was 1 ml. To quench the reactions, 200µl-aliquots were taken from the reaction mixtures after 5, 15, 30, 60, and 120 min, and NH$_4$HCO$_3$ (20mM final concentration) was added and the sample was stored at -20°C.

Table 6-2: Composition of the CaM / Mel Reaction Mixtures with the Cross-Linkers sDST and BS^3

	0x	10x	50x	100x	Final conc.
20mM HEPES (pH 7.4)	784.5µl	784.5µl	784.5µl	784.5µl	20mM
100mM CaCl$_2$**	10µl	10µl	10µl	10µl	1mM Ca^{2+}
1mg/ml Melittin**	28.5µl	28.5µl	28.5µl	28.5µl	10µM
1mg/ml Calmodulin**	167µl	167µl	167µl	167µl	10µM
sDST / BS^3 stock solution*	DMSO 10µL	10mM 10 µl	50mM 10 µl	100mM 10 µl	0, 100µM, 500µM, 1000µM

* in DMSO   ** in HEPES buffer

6.2.3.2   Cross-Linking Reactions of the Annexin A2 / P11 Complex

The homobifunctional cross-linking reagents sDST, BS^3, BS^{2}G, and DSA were used for chemical cross-linking of the ANXA2 / p11 complex. The latter three reagents were applied as 1:1 mixtures of non-deuterated and deuterated species (d_0/d_4 or d_0/d_8).

sDST

A 5-µM solution of A2t containing 1 mM CaCl$_2$ (100 mM MES buffer; 150 mM NaCl, 1 mM DTT, pH 7.5; final volume 120 µl) was incubated at room temperature for 10 min prior to addition of the homobifunctional cross-linker sDST. sDST was dissolved in DMSO (2 and 5 mM stock

solutions, prepared shortly prior to addition) and 6 µl were added to give a 20- and 50-fold molar excess over the protein concentration. For quenching the reaction, 40µl-aliquots were taken from the reaction mixtures after 15, 30, and 60 min, and NH_4HCO_3 (20 mM final concentration) was added.

Table 6-3: Composition of the ANXA2 / P11 Reaction Mixture with the Cross-Linker sDST

	20x	50x	Final conc.
A2t** 5.46 µM	110 µl	110 µl	5µM
30mM $CaCl_2$**	4 µl	4 µl	1mM Ca^{2+}
sDST* stock solution	2mM	5mM	100µM, 250µM
added volume	6 µl	6 µl	

* in DMSO   ** in MES buffer

BS³-d₀/d₄, BS²G-d₀/d₄, and DSA-d₀/d₈

The homobifunctional cross-linking reagents BS^3, BS^2G, and DSA were employed as 1:1 mixtures of their non-deuterated and deuterated form (BS^3-d_0/d_4, BS^2G-d_0/d_4, and DSA-d_0/d_8). Stock solutions (75 and 150 mM) of cross-linking reagents were prepared shortly prior to application. 6 µl were added to 594 µl of a 1.5 µM A2t solution in 20 mM HEPES, 150 mM NaCl, 1 mM DTT containing 1 mM $CaCl_2$ (pH 7.4) to result in a 50- and 100-fold molar excess over the protein concentration (600 µl final volume). To quench the reactions, 200-µl-aliquots were taken from the reaction mixtures after 30, 60 and 120 min, and NH_4HCO_3 (20 mM final concentration) was added.

Table 6-4: Composition of A2t Reaction Mixtures with Isotope-Labeled Cross-Linkers

	50x	100x	Final conc.
A2t 1.5µM in			1.5 µM
20mM HEPES (pH 7.4),	594 µl	594µl	20mM
1mM $CaCl_2$			1mM Ca^{2+}
X-linker* stock solution	7.5mM	15mM	75 µM, 100 µM
added volume	6 µl	6 µl	

* in DMSO

6.2.4   Identification of A2t Interaction Partners by Chemical Cross-Linking

The protocol for identification of A2t binding partners involves several steps. The first step comprises the biotinylation (and labeling with photoreactive cross-linker) of A2t. For testing biotinylation efficiency, A2t (1.5 µM, in 20mM HEPES buffer, pH 7.5) was reacted with 1, 2, and 4 mM aqueous solutions of the sulfhydryl-reactive labeling reagent (+)-biotinyl-iodoacetamidyl-3,6-dioxaoctane diamine (EZ-Link PEO-iodoacetyl biotin, Pierce) for 45 and 90 minutes at 37 °C. The reaction was quenched by the addition of five-fold excess dithiothreitol (DTT) over

labeling reagent concentration. The extent of biotinylation was evaluated by MALDI-TOFMS of *in-gel* tryptic digests of modified ANXA2 and subsequent search with the ExPASy *FindMod* tool and the GPMAW software, with PEO iodoacetyl biotin entered as variable modification (Δm = 414.194 u). For conducting the cross-linking experiments of A2t with its binding partners, a 4mM solution PEO iodoacetyl biotin and a reaction time of 45 minutes was chosen.

Two cross-linking strategies were developed for the identification of A2t interaction partners: In preparation A, A2t (6 ml, 1.5 µM, 20mM HEPES buffer, pH 7.5) was biotinylated (4mM PEO iodoacetyl biotin) and simultaneously labeled with the heterobifunctional amine- and photoreactive cross-linker *N*-succinimidyl-6-[4'-azido-2'-nitrophenylamino]hexanoate (SANPAH) added at a 50-fold molar excess over A2t and reacted at its amine-reactive site (45 min incubation time). In preparation B, A2t (6 ml, 1.5 µM) was only labeled with biotin (4mM solution of PEO iodoacetyl biotin). Labeling of amines was quenched by addition of Tris, and DTT (5-fold excess) was used for stopping the biotinylation reaction. In the second step, mucosal scrapings were thoroughly homogenized and washed (two times with 1 l 30 mM HEPES / 0.6 M NaCl, pH 7.5 and two times with 1 l 30 mM HEPES / 0.15 M NaCl, pH 7.5) in the presence of 1 mM $CaCl_2$ and 1 mM DTT (supernatants were discarded). *Complete EDTA-free* tablets were used for protease inhibition. Then prelabeled A2t of preparation A and B (6 ml, 1.5 µM) was added to two individual mucosal preparations (~50 ml each), respectively, and allowed to interact with potential interaction partners at ambient temperature for 20 minutes. For covalent attachment of A2t with its binding partners (third step), the mixture containing preparation A (A2t labeled with amine-reacted SANPAH) was UV-irradiated for 60 minutes. Ethylene glycol *bis*(succinimidyl succinate) (EGS) was added to sample preparation B containing A2t labeled with PEO iodoacetyl biotin only. EGS reaction was stopped with excess Tris. For both preparations, extraction of soluble (i.e. not membrane-associated) cross-linked proteins was accomplished by thorough cell disruption, addition of EGTA (10mM) and ultracentrifugation (45,000 rpm, 45 min). The supernatant containing cytosolic biotin-labeled A2t / binding partner complexes was retained. Then SDS (0.5%) was added to the pellet for disintegrating the membrane, thus releasing membrane-associated proteins. The supernatant of the subsequent centrifugation was retained. For obtaining biotin-labeled A2t / membrane-associated protein complexes the samples were purified by affinity purification on UltraLink Immobilized Monomeric Avidin (Pierce) beads. Two-times (one each for preparation A and preparation B) 450 µl avidin beads were washed with 1.5 ml 100 mM HEPES, pH 7.5 for 10 min and then centrifuged 3 min at 5500 rpm (Eppendorf MiniSpin centrifuge). The washing procedure was repeated once with 1.5 ml 100 mM glycine, pH 2.8 and three-times 1.5 ml 100 mM HEPES, pH 7.5. Avidin beads were treated with biotin (1.5 ml 2mM biotin for 30 min). Biotin was removed by washing again one time with glycine and then with 15% MeOH in 50 mM NH_4HCO_3, pH 7.5. The beads were equilibrated with 1.5 ml HEPES (*see above*). Approximately 100 µl sample were added per 100

µl beads and twice the amount HEPES (30 mM, pH 7.5) was added and the mixture was incubated for 3 hours on a shaker. Afterwards the beads were centrifuged and both supernatant and beads were retained. Three consecutive washing steps with HEPES (see procedure above) removed unbound sample. Elution of affinity-purified samples was achieved by two washing steps with 50% ACN and 0.4% TFA with incubation times of 20 minutes and subsequent centrifugation. The supernatants were dried in a vacuum concentrator. The obtained biotin-labeled proteins were separated by SDS-PAGE and gel bands were excised and tryptically digested (chapter 6.2.6.1) for subsequent analysis by MALDI-TOFMS and protein identification by peptide mass fingerprinting. In addition to the Autoflex I instrument, the samples were analyzed on an Ultraflex III MALDI-TOF/TOF instrument. Tandem mass spectrometric data was obtained for most of the samples.

6.2.5 Polyacrylamide Gel Electrophoresis

One-dimensional polyacrylamide gel electrophoresis (PAGE) was carried out using the vertical Mini-PROTEAN 3 electrophoresis system with POWERPAC 300 power supply (Bio-Rad, München, Germany).

SDS gels, consisting of a stacking (5%, pH 6.8) and a resolving gel (10, 12, and 15%, pH 8.8), were prepared with the ingredients listed in **Table 6-5** (gel size: 8.0 cm x 7.3 cm x 0.075 cm). First, the resolving gel was prepared by combining 40% acrylamide / bisacrylamide (37.5:1, 2.6% C) solution, 1.5 M Tris-HCl (pH 8.8), sodium dodecyl sulfate (SDS), and H_2O. Addition of ammonium persulfate (APS) and *N,N,N',N'*-tetramethylethylenediamine (TEMED) initiated the polymerization and the mixture was immediately poured between the glass plates and overlayed with isopropanol. After 30 minutes, the gel was polymerized and the isopropanol layer was removed (rinse with water to remove residual isopropanol). The stacking gel was prepared and added on top of the resolving gel. A 10-well square-bottom comb was inserted into the stacking gel for forming the gel pockets.

Table 6-5: Preparation of Polyacrylamide Gels.

	Stacking Gel 5%	Resolving Gel X%
40% Acrylamide / Bis	640 µl	$0.25 \cdot X\% = x$ ml
0.5M Tris-HCl, pH 6.8	1.25 ml	---
1.5M Tris-HCl, pH 8.8	---	2.5 ml
10% SDS (w/v)	50 µl	100 µl
H_2O	3.025 ml	7.35 - x ml
TEMED	5 µl	5 µl
10% APS (w/v)	25 µl	50 µl
Total Volume	5 ml	10 ml

Native 8% (pH 8.8) gels were cast as continuous gels, i.e., without stacking gel, and SDS was replaced by H_2O.

For native PAGE and SDS-PAGE different sample preparation conditions, running buffers, and electrophoresis conditions were employed as described below.

Following electrophoretic separation, gels were stained for 60 min with a staining solution containing 0.1% (w/v) Coomassie Brilliant Blue R-250, 50% (v/v) methanol, and 10% (v/v) glacial acetic acid. After 60 minutes the staining solution was replaced by the destaining solution (25% (v/v) methanol, 10% (v/v) glacial acetic acid) and the gels were destained until protein bands became visible and background stain was sufficiently removed. Gels were air-dried between cellophane sheets for documentation.

6.2.5.1 Native Polyacrylamide Gel Electrophoresis

Native PAGE of the intact, non-cross-linked as well as cross-linked A2t was performed under non-denaturing conditions in the absence of SDS and reducing agents employing 8% continuous gels. Native sample buffer (containing 62.5 mM Tris-HCl, pH 6.8, 40% glycerol (Bio-Rad)) without SDS and 2-mercaptoethanol was used for sample preparation in 1:1 ratio without heating and the samples were filled into the sample wells. Detergent-free, 5x concentrated running buffer (pH 8.3) containing 0.96 M glycine and 0.125 M Tris was diluted 1:5 (v/v). Electrophoretic separation was conducted at 100V for ~4 hours at 4°C. Immunoglobulin A (300 kDa), lactate dehydrogenase (140 kDa), phosphorylase b (97 kDa), bovine serum albumine (67 kDa), carbonic anhydrase (29 kDa), and cytochrome c (12 kDa) were used as molecular weight standards.

6.2.5.2 SDS - Polyacrylamide Gel Electrophoresis

One-dimensional SDS-PAGE was conducted as described by Laemmli [Laemmli, 1970]. Samples were mixed 1:1 (v/v) with Laemmli sample buffer containing 5% (v/v) 2-mercaptoethanol and heated for 5 min at 95°C. For separation of the CaM / melittin cross-linking reaction mixtures, 5% stacking / 15% resolving gels were prepared. For separation of the ANXA2 / p11 cross-linking reaction mixtures, 5% stacking and 10, 12, or 15% resolving gels were employed. Gels for monitoring the protein purification process and gels for subsequent Western Blot analysis were 5% stacking / 12% resolving gels. Running buffer (prepared from 10x concentrated running buffer containing 0.25 M Tris, 1.92 M glycine, and 1% (*v/v*) SDS) (Roth), was used for denaturing SDS-PAGE. Electrophoretic separation was conducted at room temperature for about 90-120 min at 140V, depending on the porosity of the gels employed.

Gels intended for Western Blotting and post-translational modification analysis were not stained with Coomassie Brilliant Blue, but were immediately processed for further use as

described in the respective chapters. Unless stated otherwise, Precision Plus Unstained molecular weight standard from Bio-Rad was used.

6.2.6   Enzymatic Proteolysis

Following separation of protein mixtures by one-dimensional polyacrylamide gel electrophoresis, the protein bands of interest were excised and *in-gel* digested. The gel bands were cut into cubes of about 1 mm^3 and the Coomassie stain was removed by three consecutive washing steps employing 150 µl of a 50% (v/v) acetonitrile (ACN) / 50% (v/v) H$_2$O solution. The gel pieces were incubated in the washing solution for 10 minutes on a shaker and the supernatants were discarded. In the next step, 80 µl ACN were added, the liquid was removed after five minutes, and was replaced by 80 µl of a 100 mM NH$_4$HCO$_3$ solution. Again, 80 µl ACN were added without removing the NH$_4$HCO$_3$ solution and the gel cubes were incubated for another ten minutes before the supernatant was discarded. In the following vacuum centrifugation step, the gel pieces were dried within 30 minutes. For reduction of disulfide bonds, 80 µl of a 10 mM dithiothreitol (DTT) / 100 mM NH$_4$HCO$_3$ solution were added and incubated at 56°C for 45 minutes. For alkylation of the reduced cysteines, the DTT solution was discarded and replaced by 80 µl of a 55mM iodoacetamide / 100 mM NH$_4$HCO$_3$ solution and incubated at ambient temperature in the dark for 30 minutes. For removal of the reagents, the gel pieces were washed by repeating the washing procedure described above and dried in the vacuum concentrator. Diverse proteases were used for different applications. Trypsin solution was prepared from 2µl-aliquots containing 1 µg trypsin in 1mM HCl. The aliquot was diluted with 18 µl of a 50 mM NH$_4$HCO$_3$ solution, resulting in a 50 ng/µl trypsin solution. Endoproteinase AspN solution was prepared by H$_2$O to result in the desired concentrations (see respective descriptions). LysC, GluC, and chymotrypsin were used at 50 ng/µl. For enzymatic proteolysis, between 5 and 10 µl of enzyme solution were added to the dry gel pieces, depending on the total volume of the gel pieces and the original Coomassie staining intensity of the gel bands. After the protease solution was fully absorbed, the gel pieces were covered by a 50 mM NH$_4$HCO$_3$ solution. The gel pieces were incubated at 37°C for 16 hours unless stated otherwise. Upon termination of the digestion by adding 50 µl ACN / H$_2$O / formic acid (47.5% (v/v) / 47.5 % (v/v) / 5% (v/v)) and retaining the supernatant, the peptides were extracted from the gel pieces by incubation at room temperature for 10 minutes with fresh solution for two more times. The individual supernatants were combined, concentrated in the vacuum concentrator to a volume of about 10-15 µl, and stored at –20°C for subsequent mass spectrometric analysis.

6.2.6.1   Enzymatic Proteolysis of Proteins for Peptide Mass Fingerprint Analysis

For characterization of amino acid sequences of CaM and melittin *in-gel* digestions were performed with trypsin (50 ng/µl, 37°C, over-night), trypsin / AspN mixture (50 ng/µl each, 37°C, over-night), LysC / AspN mixture (50 ng/µl each, 37°C, over-night), and AspN (40 ng/µl, 37°C, over-night), respectively.

ANXA2 and p11 amino acid sequence coverages were obtained by analysis of digests with trypsin (50 ng/µl, 37°C, over-night) and AspN (30 or 60 ng/µl, 37°C, over-night). Chymotrypsin (50 ng/µl, 37°C, 2hrs) was additionally employed for *in-gel* digestion of p11 for obtaining better sequence coverage. In addition to trypsin, GluC (50 ng/µl, 37°C, 16 hrs) was used for characterization of the N-terminally truncated annexin A2 species.

Co-purified proteins from A2t purification were digested with trypsin (50 ng/µl, 37°C, over-night), as were all the proteins of A2t binding partners (chapter 6.2.4).

6.2.6.2   Enzymatic Proteolysis of Calmodulin / Melittin Cross-Linking Reaction Mixtures

Trypsin alone (50 ng/µL) was used for all cross-linking reaction mixtures, whereas endoproteinase AspN (40 ng/µL), trypsin/AspN (both 100 ng/µl), and LysC/AspN (both 100 ng/µl) were used as additional digestion enzymes of the CaM / Mel complex cross-linked with EDC/sNHS. When employing a combination of two proteases, only half the volume of each was added to the gel bands to result in a two-fold dilution (i.e. 50 ng/µl) of the initial concentration.

6.2.6.3   Enzymatic Proteolysis of A2t Cross-Linking Reaction Mixtures

Trypsin (50 ng/µL) was used for all cross-linking reaction mixtures, whereas endoproteinase AspN (50 ng/µL) was exclusively used for the sDST cross-linking reaction mixtures. The digests were incubated at 37°C for 16 hrs.

6.2.7   MALDI-TOF Mass Spectrometry

MALDI-TOFMS was performed on a Voyager DE™ RP Biospectrometry™ Workstation (Applied Biosystems) and on an Autoflex I instrument (Bruker Daltonik) both equipped with a nitrogen laser (337 nm). Both instruments possess delayed-extraction technology and were operated in positive ionization mode. Furthermore, both mass spectrometers possess a linear and a reflector detector.

For analyses with the Voyager instrument 96-well stainless steel MALDI sample plates were used. An MTP 384 massive target T was the target plate for the Autoflex I instrument.

Data acquisition and data processing were performed using the Voyager software version 5.1 and the Data Explorer software version 4.0 (Applied Biosystems), and the Flex Control 2.2.19.0 and Flex Analysis 2.2 software (Bruker Daltonik).

6.2.7.1 Sample Preparation for MALDI-TOFMS

Prior to MALDI-TOF mass spectrometric analysis the target plate onto which the samples are to be deposited needs to be thoroughly cleaned. This is achieved by rinsing with different solvents and gentle rubbing of the target surface.

The samples were desalted prior to analysis employing either C-18 or C-4 ZipTips (Millipore) for low-volume peptide and protein samples, respectively. For larger volumes, microcon centrifugal filter devices (Millipore) with appropriate cut-off were employed. The matrices were prepared by preparing a saturated solution in 50-70 % (v/v) acetonitrile and 0.1 % (v/v) TFA or FA. The supernatant was several-fold diluted until a thinly and evenly dispersed matrix layer was observed upon deposition on the target. α-Cyano-4-hydroxy cinnamic acid (Bruker, (Sigma)) was employed for peptide analyses. Sinapinic acid (Sigma) was used for linear MALDITOFMS of the intact CaM/Mel complex. For linear MALDI-TOFMS of intact ANXA2 the following matrices were employed: 6-aza-2-thiothymine (ATT), 2',5'-dihydroxyacetophenone , 2,5-dihydroxybenzoic acid (DHB), 3,5-dimethoxy-4-hydroxy cinnamic acid (sinapinic acid), and 4-hydroxyazobenzene-2-carboxylic acid (HABA).

Matrix and sample solutions were deposited on the MALDI sample plate using the dried-droplet method.

6.2.7.2 Voyager DE RP

Linear mode MALDI-TOFMS of the non-digested CaM / Mel cross-linking reaction mixtures was performed for monitoring the extent of chemical cross-linking. Measurements were conducted in positive ionization mode using sinapinic acid as the matrix.

The matrix was prepared in 30% (v/v) acetonitrile / 70% (v/v) H_2O / and 0.1% (v/v) trifluoro acetic acid. Samples were desalted employing Microcon YM-10 filters (Millipore) and prepared using the dried droplet method. 100 shots were added to one spectrum in the mass range m/z 2000 to 42000. The instrument was calibrated using cytochrome c ($[M+H]^+_{average}$ at m/z 12361.6) and myoglobin ($[M+H]^+_{average}$ at m/z 16952.6).

For sequence coverages of CaM, melittin, and ANXA2, the instrument was run in positive reflector mode with typically 200 shots summed to one spectrum. Calibration was performed with angiotensin II ($[M+H]^+_{mono}$ at m/z 1046.54), angiotensin I ($[M+H]^+_{mono}$ at m/z 1296.68), substance P ($[M+H]^+_{mono}$ at m/z 1347.74), and somatostatin (red) ($[M+H]^+_{mono}$ at m/z 1637.72).

6.2.7.3 Autoflex I

For protein analysis, measurements were performed in linear positive ionization mode using sinapinic acid, 2,5-dihydroxybenzoic acid (DHB), 2',5'-dihyroxyacetophenone, 4-hydroxyazobenzene-2-carboxylic acid (HABA), and 6-aza-2-thiothymine (ATT) as matrices.

Positive ionization and reflectron mode were employed for MALDI-TOFMS measurements of peptide mixtures using α-cyano-4-hydroxy-cinnamic acid as matrix. All matrices were prepared in 50% (v/v) acetonitrile / 50% (v/v) H_2O / and 0.1% (v/v) formic acid. Samples were desalted using C4-ZipTips for proteins and C18-ZipTips for peptides (Millipore), respectively, and prepared using the dried droplet method.

For peptide analysis, 200 laser shots were added to one spectrum typically in the *m/z* ranges 800 to 4,000 or *m/z* 500 to 4,000. Peptide Calibration Standard (Bruker Daltonik) (angiotensin II ($[M+H]^+_{mono}$ at *m/z* 1046.54), angiotensin I ($[M+H]^+_{mono}$ at *m/z* 1296.68), substance P ($[M+H]^+_{mono}$ at *m/z* 1347.74), bombesin ($[M+H]^+_{mono}$ at *m/z* 1619.82), ACTH clip 1-17 ($[M+H]^+_{mono}$ at *m/z* 2093.09), ACTH clip 18-39 ($[M+H]^+_{mono}$ at *m/z* 2465.20), and somatostatin ($[M+H]^+_{mono}$ at *m/z* 3147.47)) was used for external calibration of the mass spectra in the reflectron mode. Additionally, internal calibration of mass spectra was performed with autolytic peptide signals of trypsin. When analyzing AspN-digested samples of ANXA2 and p11, signals of ANXA2 and p11 peptides generated from AspN digestion that were frequently observed in the spectra, were used for internal calibration. For external mass spectra calibration in the linear mode, cytochrome c ($[M+H]^+_{average}$ at *m/z* 12362), carbonic anhydrase ($[M+H]^+_{average}$ at *m/z* 29025), and bovine serum albumin ($[M+H]^+_{average}$ at *m/z* 66431) were used. 1000 laser shots were accumulated to one spectrum in the *m/z* range between 8,000 and 110,000.

Ultraflex III: Experiments were conducted at Bruker Daltonik, Bremen.

6.2.8   Nano-High Performance Liquid Chromatography / Nano-Electrospray Ionization-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

6.2.8.1   *Mass Spectrometric Analysis of the Calmodulin / Melittin Complex*

The peptide mixtures from enzymatic digests were separated by nano-HPLC. Nano-HPLC was carried out on an Ultimate™ Nano-LC system (LC Packings) equipped with a Switchos II column switching module and a Famos™ Micro Autosampler with a 5-µl sample loop. Samples were injected by the autosampler and concentrated on a trapping column (PepMap, C18, 300 µm * 1 mm, 5 µm, 100 Å, LC Packings) with water containing 0.1% formic acid (v/v) at flow rates of 20 µl/min. After two minutes, the peptides were eluted onto the separation column (PepMap, C18, 75 µm * 150 mm, 3 µm, 100 Å, LC Packings), which had been equilibrated with 95% A (A being H_2O + 0.1% (v/v) formic acid). Peptides were separated using the following gradient: 0-30 min: 5-50% B, 30-31 min: 50-95% B, 31-35 min: 95% B (B being acetonitrile + 0.1% (v/v) formic acid) at flow rates of 200 nl/min and detected by their UV absorptions at 214 and 280 nm.

The nano-HPLC system was coupled *on-line* to an Apex II FTICR mass spectrometer equipped with a 7 Tesla superconducting magnet (Bruker Daltonics) and a nano-electrospray ionization source (Agilent Technologies). For nano-ESIMS, coated fused-silica PicoTips (tip IDs 8 µm and 15 µm, New Objective) were applied. The capillary voltage was set to -1400 V. Mass spectral data were acquired over a range of *m/z* 400-2000, four scans were accumulated per spectrum, and 400 spectra were recorded for each LC/MS run. MS data acquisition was initialized with a trigger signal from the HPLC system 5 min after initiation of the LC gradient. Data were acquired over 34.5 min. Calibration of the instrument was performed with CID fragments (capillary exit voltage 200 V) of the LHRH peptide. Data acquisition and data processing were performed using the XMASS software, versions 5.0.10 and 6.0 (Bruker Daltonics). Processing of the raw data was performed using the '*Projection*' tool in the XMASS software [Bruker BioAPEX User's Manual, 1996].

6.2.8.2   *Mass Spectrometric Analysis of the ANXA2 / P11 Complex*

The peptide mixtures from enzymatic digests were separated by the nano-HPLC system as described above. The only differences were that after three minutes peptides were eluted onto the separation column and that data acquisition and data processing were performed using XMASS version 7.02. [2]

6.2.8.3   *Tandem Mass Spectrometric Analysis of the ANXA2 / P11 Complex*

In-gel digests of cross-linking reaction mixtures from heterodimeric and heterotetrameric complexes between ANXA2 and p11 were additionally analyzed by MS/MS experiments. The peptide mixtures resulting from the *in-gel* digests were separated by C18-RP-chromatography on a nano-HPLC system (Ultimate 3000, Dionex; pre-column: PepMap C18, 300µm * 5 mm, 3 µm, 100 Å, Dionex; separation column: PepMap, C18, 75 µm * 150 mm, 3 µm, 100 Å, Dionex; solvents A: 5% (v/v) acetonitrile, 0.1% (v/v) formic acid in water, B: 80% (v/v) acetonitrile, 0.08% (v/v) formic acid in water) using a gradient from 0 to 60 % B in 90 minutes followed by isocratic elution with 90% B for 3 minutes. An LTQ-FT mass spectrometer (Thermo Electron) with a 7 Tesla magnet equipped with a nano-ESI source (Proxeon Biosystems; emitter: distal coated PicoTips, tip i.d. 15 µm, New Objective) was *on-line* coupled to the nano-HPLC system. MS data were acquired over 100 min in data-dependent MS[2] mode: each high-resolution full scan (*m/z* 300–2,000, resolution at *m/z* 400 was set to 100,000) in the ICR cell was followed by 10 product ion scans in the linear trap of the 10 most intense signals in the full-scan mass

---

[2] Measurements on the ApexII instrument were conducted by Christian Ihling and Stefan Kalkhof.

spectrum (isolation window 3 u). Dynamic exclusion (exclusion duration 20 s, exclusion window ± 5 ppm) was enabled in order to allow detection of less abundant ions. [3]

6.2.9 Processing of Mass Spectra

6.2.9.1 MALDI-TOFMS Data

Labeling of the mass spectra recorded on the Voyager (Applied Biosystems) instrument in the reflector mode was performed by the *Deisotoping* tool in the Data Explorer software (v. 4.0, Applied Biosystems). The obtained monoisotopic masses ([M+H]$^+$) were manually evaluated by setting peak detection thresholds for defined detection ranges based on signal intensities. Linear MALDI-TOF mass spectra were labeled with the *Peak Label* tool in the Data Explorer software.

Monoisotopic masses ([M+H]$^+$) in the mass spectra recorded on the Autoflex I (Bruker Daltonik) in the reflector mode were labeled manually. Linear MALDI-TOF mass spectra were as well evaluated manually.

6.2.9.2 ESI-FTICRMS Data

ESI-FTICR (Apex II) mass spectra were processed with the XMass (v. 7.0.2 and 7.0.3, Bruker Daltonics) software. The single spectra were projected into one final mass spectrum using the *Projection* tool in the XMass software. The mass spectra were deconvoluted and monoisotopic masses ([M+H]$^+$) were manually labeled.

6.2.9.3 MS/MS Data

Data analysis was performed by dividing the chromatograms of the total ion current into five-minutes-fractions and the obtained average mass spectra were deconvoluted using the Mascot Distiller software v. 2.0 (Matrix Science). Product ion mass spectra of identified cross-linked peptides were also processed with Mascot Distiller (v. 2.0, Matrix Science).

6.2.10 Peptide Mass Fingerprint Analysis

The Mascot (www.matrixscience.com) search tool *Peptide Mass Fingerprinting* was employed for the identification of unknown proteins obtained during A2t purification. A set of experimentally obtained peptide masses is compared to theoretical peptide masses calculated from protein sequences deposited in protein databases like e.g. the NCBI (National Center for Biotechnology Information, www.ncbi.nlm.nih.gov) database. Information on the taxonomy, the

---

[3] Measurements on the LTQ-FT instrument were performed by Christoph Stingl.

employed protease, the number of missed cleavages, and on fixed (e.g. carbamidomethylation, N-terminal acetylation) and variable (e.g. methionine oxidation) modifications are provided by the user. Furthermore, the error window for experimental peptide masses needs to be defined. Then the entered MS peaklist is searched for matches and proteins putatively matching the entered masses are displayed. The probability-based MOWSE score given for an identified protein is a criterion whether the result is significant or not.

6.2.11 Data Analysis

Cross-linked products were identified using IsoFind, the GPMAW (General Protein Mass Analysis for Windows) software, versions 5.12beta3, 6.00, and 7.01 (Lighthouse Data) (available at: http//welcome.to/gpmaw), the ExPASy Proteomics tools in the Swiss-Prot Database (available at: www.expasy.org), ASAP (Automatic Spectrum Assignment Program) v. 1.09 and MS2Assign.

6.2.11.1 IsoFind

The *in-house* developed tool 'IsoFind' (developed by Tibor Kohajda and Stefan Kalkhof) was employed for searching for distinct 4.025 or 8.05 amu spacing between two signals from a given MS peaklist. The macro is run in Excel and applied to MS peaklists containing *m/z* values in one column and either signal intensity or peak area values in the second column. From this, the macro additionally calculates the signal ratios based on either signal intensity or peak area.

6.2.11.2 General Protein Mass Analysis for Windows (GPMAW)

This software calculates theoretical cross-linked products (inter- and intramolecular cross-linked products, peptides modified by a partially hydrolyzed cross-linking reagent) and peptides for one or two given protein sequences. Both the protease and cross-linking reagent employed are defined by the user as well as the maximum number of missed cleavages. The experimentally obtained masses from mass spectrometric analysis are compared to the theoretic values and are assigned to cross-linked products or peptides. Maximum mass deviations for correct assignment were 10 ppm for ESI-FTICRMS data and (at maximum) 100 ppm for MALDI-TOFMS data

In case of doubt, cross-links assigned by the software were not considered as such when discrimination between a peptide and a cross-linked product or between two different cross-linked products, coincidently having the same mass, was ambiguous. In case of the CaM / melittin complex, cleavages by proteases at modified amino acids, such as the trimethylated K115 in CaM as well as amino acids modified by cross-linking reagents, were excluded. Both the N-terminus of CaM and the C-terminus of melittin were excluded from possible cross-linking

by EDC since the former is acetylated and the latter amidated. In case of the ANXA2 / p11 complex, cleavages by proteases at amino acids modified by cross-linking reagents, were allowed. The N-terminus of ANXA2 is acetylated and thus unavailable for cross-linking.

6.2.11.3 ExPASy Proteomics Tools

The *FindPept* and *FindMod* tools of the Swiss-Prot database (www.expasy.org) were used for comparison of experimentally obtained *m/z* values from MS analysis of an enzymatically digested protein with the theoretical peptide masses calculated for a specified amino acid sequence. The *FindPept* tool was also used for searching for intrapeptide cross-links and peptides with a hydrolyzed cross-linker by defining the cross-linking reagent as variable modification. Data evaluation of the biotinylated protein was also performed using the *FindPept* and *FindMod* tools. The program also accounts for protease autolysis and contaminant keratin peptides.

6.2.11.4 Automated Spectrum Assignment Program (ASAP)

The ASAP tool (http://roswell.ca.sandia.gov/~mmyoung/asap.html) was used for assigning *m/z* values obtained from MS analysis of enzymatically digested cross-linked proteins. ASAP assigns cross-linked products and peptides based on comparison to theoretical cross-linked products and peptides for a given protein. Information on cross-linker, protease, and amino acid modifications has to be provided. Unfortunately, ASAP works only for single proteins. Nevertheless, this program is useful as it, in contrast to GPMAW, considers variable modifications.

6.2.11.5 MS2 Assign

MS2 Assign (http://roswell.ca.sandia.gov/~mmyoung/ms2assign.html) was used for assigning *m/z* values obtained from MS fragmentation data of cross-linked peptides. MS2Assign calculates a theoretical fragment library for a (cross-linked) peptide or a cross-linked pair of peptides. Peptide sequences, crosslinker (if applicable), and amino acid modifications have to be provided.

6.2.12   Computational Protein-Protein Docking

6.2.12.1 Protein-Peptide Docking of the CaM / Melittin Complex with Xplor-NIH

Structures of the CaM / melittin complex were calculated by conjoined rigid body/torsion angle simulated annealing [Clore, 2000, Schwieters & Clore, 2001, Clore & Bewley, 2002] using the molecular structure determination package Xplor-NIH [Schwieters *et al.*, 2003] (available *on-line*

at: http://nmr.cit.nih.gov/xplor-nih) and structures were viewed using the VMD-XPLOR visualization package [Schwieters & Clore, 2001] (available at: http://vmd-xplor.cit.nih.gov). Distance restraints derived from cross-linking data were represented by empirical $<r^{-6}>^{-1/6}$ averages. This ensures that at least one of the distances in the restraint lies within the specified target range, but no penalty ensues if the remaining distances within the restraint are much longer than the target distance. The distances were classified into three ranges, ≤ 5 Å, ≤ 8.5 Å and ≤ 13.4 Å, corresponding to cross-links obtained using EDC, sDST and BS[3], respectively. The starting coordinates employed for CaM and melittin were derived from the 2.2 Å resolution crystal structure of the CaM / smMLCK peptide complex (PDB accession code 1CDL [Meador *et al.*, 1992) and the 2 Å resolution crystal structure of helical melittin (PDB accession code 2MLT [Eisenberg *et al.*, 1980]), respectively. The N- (residues 5-70) and C-terminal (residues 81-146) halves of CaM, and the N- (residues 1-9) and C (residues 13-26) terminal halves of melittin were treated as rigid bodies. Residues 71-80 of calmodulin, residues 10-12 of melittin (i.e., the location of the kink between the two helical segments) and all interfacial side chains were given full torsional degrees of freedom. The non-bonded term in the target function comprised a quartic van-der-Waals repulsion term [Nilges *et al.*, 1988] and a torsion angle database potential of mean force derived from high-resolution crystal structures [Clore & Kuszewski, 2002]. The latter ensures that the side chains torsion angles lie within energetically allowed rotamer ranges. A qualitative interpretation of the cross-linking data indicated that melittin could bind CaM in two opposing orientations. Consequently, both orientations were refined simultaneously by including two complete sets of non-overlapping coordinates (i.e., no interactions were allowed between the two sets). [4]

6.2.12.2 Protein-Protein Docking of A2t with Rosetta

The X-ray structures of N-terminally truncated human ANXA2 (PDB entry 1W7B [Rosengarth *et al.*, 2004]) and the heterotetramer comprised of one p11 dimer (*Homo sapiens*) and two synthetic ANXA2 N-terminal peptides (PDB entry 1BT6 [Réty *et al.*, 1999]) were used as starting structures for the docking procedure with Rosetta [Gray *et al.*, 2003, Daily *et al.*, 2005]. In a first low-resolution docking step 20,000 models were generated by randomizing both docking partners. The models were filtered using the four distance restraints listed in **Table 4-5**. The remaining 125 structures were clustered twice with cluster thresholds of 5 and 9 Å rmsd (root mean square deviation), respectively. The structures representing the centers of clusters with more than four models were used as starting structures for the following high-resolution docking run, allowing a perturbation of 5 and 9 Å, respectively. The resulting structures were filtered again (**Table 4-5**) and clustered with an rmsd threshold of 5 Å. The cluster center

---

[4] Computational modeling of the calmodulin / melittin complex was performed by Dr. G. Marius Clore

structures of the top nine clusters with more than four models, together with two additional structures, retrieved from filtering for NZ-NZ distances of lysine side chains, were used for generating models of the symmetric A2t complex. A second ANXA2 monomer was added to the $(p11)_2$ / ANXA2 complex by superimposition, utilizing the inherent symmetry of the p11 dimer. This superimposition, which was done with MAMMOTH [Olmea *et al.*, 2002], led to steric clashes for three out of eleven structures. For the remaining eight structures, the missing eight amino acids between the N-terminal annexin peptide and N-terminally truncated ANXA2 as well as the four C-terminal amino acids of p11 were modeled based on secondary structure predictions using JUFO (www.meilerlab.org), PSIpred (http://bioinf.cs.ucl.ac.uk/psipred/), and SAM (http://www.cse.ucsc.edu/research/compbio/sam.html). Amino acid substitutions between porcine and human ANXA2 as well as between porcine ANXA2 and synthetic N-terminal ANXA2 peptide were conducted before the A2t structures were subjected to repacking and side-chain relaxation. With the objective of obtaining a model of the octameric ANXA2 / p11 complex, we created 2000 models of a p11 homotetramer by randomized docking. The starting structure for the subsequent dock perturbation run was chosen restricting the N- and C-termini of the two respective dimers to be in close proximity, thus ensuring that the opposite site that was found to be the ANXA2 binding site remained unoccupied. From the 2000 model structures after dock perturbation, one structure was picked, complying to the premise that the N- and C-termini of the two respective dimers are equidistant and at the same time close to each other, thus obtaining a compact and symmetric homotetramer. The models of the heterooctamer structures between ANXA2 and p11 were obtained from the overlay of the A2t structure based on the geometry of the p11 homotetramer. The dimensions of the created octamer structures were determined using the *Tabling* function in the AMoRe software [Navaza, 1994]. [5]

6.2.12.3   Determination of Solvent Accessibilities of Amino Acid Sidechains

Partially hydrolyzed cross-linkers, which are attached to lysine residues, provide valuable information on the solvent accessibilities of the respective amino acid side chains in a protein. The modified lysine residues were used for creating surface topology maps of ANXA2 and p11. The findings were validated by the GETAREA 1.1 software (www.scsb.utmb.edu/getarea, [Fraczkiewicz & Braun, 1998]) that calculates the solvent accessibilities of amino acid residues and assigns amino acid residues that display more than 50% solvent accessibility to be exposed on the surface of a protein.

---

[5] The expert assistance of Stefan Kalkhof, who significantly contributed to computing the models of the complex between annexin A2 and p11, is gratefully acknowledged.

References

Abagyan, R., Totrov, M., and Kuznetsov, D. (1194) ICM - A new method for protein modeling and design, Applications to docking and structure prediction from the distorted native conformation. *J. Comp. Chem*. **15**, 488-506

Anand, G. S., Law, D., Mandell, J. G., Snead, A. N., Tsigelny, I., Taylor, S. S., Ten Eyck, L. F., and Komives, E. A. (2003) Identification of the protein kinase A regulatory R-I alpha-catalytic subunit interface from amide H/H-2 exchange and protein docking. *Proc. Natl. Acad. Sci. USA* **100**, 13264-13269

Babu, Y. S., Sack, J. S., Greenhough, T. J., Bugg, C. E., Means, A. R., and Cook, W. J. (1985) 3-Dimensional structure of calmodulin. *Nature* **315**, 37-40

Babu, Y. S., Bugg, C. E., and Cook, W. J. (1988) Structure of calmodulin refined at 2.2 Å resolution. *J. Mol. Biol*. **204**, 191-204

Back, J. W., Jong, L. D., O.Muijsers A., and Koster, C. G. (2003) Chemical cross-linking and mass spectrometry for protein structural modeling. *J. Mol. Biol*. **331**, 301-313

Bähler, M. and Rhoads, A. (2002) Calmodulin signaling via the IQ motif. *FEBS Lett*. **513**, 107-113

Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., and Bax, A. (1992) Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy, The central helix is flexible. *Biochemistry* **31**, 5269-5278

Barth, A., Martin, S. R., and Bayley P. M. (1998) Specifity and Symmetry in the Interaction of Calmodulin Domains with the Skeletal Muscle Myosin Light Chain Kinase Target Sequence. *J. Biol. Chem*. **273**, 2174-2183

Baumeister, W. and Steven, A. C. (2000) Macromolecular electron microscopy in the era of structural genomics. *Trends Biochem. Sci*. **25**, 12, 624-631

Baumruk V., Pancoska P., and Keiderling, T. A. (1996) Predictions of secondary structure using statistical analysis of electronic and vibrational circular dichroism and Fourier Transform Infrared spectra of proteins in H_2O. *J.Mol.Biol*. **259**, 4, 774-791

Becker T, Weber K, and Johnsson N. (1990) Protein-protein recognition via short amphiphilic helices, a mutational analysis of the binding site of annexin II for p11. *EMBO J*. **9**, 4207-4213

Biemann, K. *In:* McCloskey, J.A., Editor, Methods in Enzymology **193**, Academic, San Diego (1990), pp. 886–887.

Biener, Y., Feinstein, R., Mayak, M., Kaburagi, Y., Kadowaki, T., and Zick, Y. (1996) Annexin II is a novel player in insulin signal transduction. *J. Biol. Chem*. **271**, 29489-29496

BioAPEX User's Manual (1996) v 1.1, Bruker Daltonics, Billerica, MA, USA

Brokx, R. D., Lopez, M. M., Vogel, H. J., and Makhatadze, G. I. (2001) Energetics of peptide binding by calmodulin reveals different modes of binding. *J. Biol. Chem*. **276**, 14083-14091

Brownstein, C., Deora, A. B., Jacovina, A., Weintraub, R., Gertler, M., Khan, K. M. F., Falcone, D. J., and Hajjar, K. A.. (2004) Annexin II mediates plasminogen-dependent matrix invasion by human monocytes, enhanced expression by macrophages. *Blood* **103**, 317-324

Burger, A., Berendes, R., Liemann, S., Benz, J., Hofmann, A., Göttig, P., Huber, R., Gerke, V., Thiel, C., Römisch, J., and Weber, K. (1996) The crystal structure and ion channel activity of human annexin II, a peripheral membrane protein. *J. Mol. Biol*. **257**: 839-847

Cai, M., Williams, D. C., Wang, G., Lee, B. R., Peterkofsky, A. and Clore, G. M. (2003) Solution structure of the phosphoryl transfer complex between the signal-transducing protein IIAGlucose and the cytoplasmic domain of the glucose transporter IICBGlucose of the Escherichia coli glucose phosphotransferase system. *J. Biol. Chem*. **278**, 25191-25206

Carafoli, E. (2002) Calcium signaling, A tale for all seasons, *Proc. Natl. Acad. Sci. USA* **99**, 1115-1122

Clore, G. M. and Bewley, C. A. (2002) Using conjoined rigid body/torsion angle simulated annealing to determine relative orientation of covalently linked protein domains from dipolar couplings. *J. Magn. Reson*. **154**, 329-335

Clore, G. M. (2000) Accurate and rapid docking of protein-protein complexes on the basis of intermolecular nuclear Overhauser enhancement data and dipolar couplings by rigid body minimization. *Proc. Natl. Acad. Sci. U.S.A*. **97**, 9021-9025

Clore, G. M. and Kuszewski, J. (2002) χ1 rotamer populations and angles of mobile surface side chains are accurately predicted by a torsion angle database potential of mean force, *J. Am. Chem. Soc*. **124**, 2866-2867

Clore, G. M., Bax, A., Ikura, M., and Gronenborn, A. M. (1993) Structure of calmodulin-target peptide complexes. *Curr. Opin. Struct. Biol*. **3**, 838-845

Comisarov, M. B. and Marshall A. G. (1974) Frequency-Sweep FT-ICR spectroscopy. *Chem. Phys. Lett*. **26**, 489-490

Comisarov, M. B. and Marshall A. G. (1996) The early development of Fourier Transform Ion Cyclotron Resonance (FT-ICR) mass spectroscopy. *J. Mass Spectrom*. **31**, 581-585

Comte, M., Maulet, Y., and Cox, A. J. (1983) Ca^{2+}-dependent high-affinity complex formation between calmodulin and melittin. *Biochem. J*. **209**, 269-272

Cornilescu, G., Lee, B. R., Cornilescu, C. C., Wang, G., Peterkofsky, A. and Clore, G. M. (2002) Solution structure of the phosphoryl transfer complex between the cytoplasmic A domain of the mannitol transporter IIGlucose and HPr of the Escherichia coli phosphotransferase system. *J. Biol. Chem*. **277**, 42289-42298

Cox, J. A., Comte, M., Fitton, J. E., and DeGrado, W. F. (1985) The interaction of calmodulin with amphiphilic peptides. *J. Biol. Chem*. **260**, 2527-2534

Crivici, A. and Ikura, M. (1995) Molecular and structural basis of target recognition by calmodulin. *Annu. Rev. Biophys. Biomol. Struct*. **24**, 85-116

Cuppoletti, J., Blumenthal, K. E., and Malinowska, D. H. (1989) Melittin inhibition of the gastric (H$^+$/K$^+$) ATPase and photoaffinity-labeling with [I-125] azidosalicyl melittin. *Arch. Biochem. Biophys*. **275**, 263-270

Cuppoletti, J. (1990) [I-125] azidosalicyl melittin binding domains – evidence for a polypeptide receptor on the gastric (H$^+$/K$^+$) ATPase. *Arch. Biochem. Biophys*. **278**, 409-415

Daily, M. D., Masica, D., Sivasubramanian, A., Somarouthu, S., and Gray, J. J. (2005) Capri Rounds 3-5 Reveal Promising Successes and Future Challenges for RosettaDock. *Proteins* **60**, 181-186

Danielsen, E. M., van Deurs, B., and Hansen, G. H. (2003) „Nonclassical" secretion of annexin A2 to the lumenal side of the enterocyte brush border membrane. *Biochemistry* **42**, 14670-14676

Delmer, D. P., and Potikha, T. S. (1997) Structures and functions of annexins in plants. *Cell. Mol. Life Sci*. **53**, 546-553

Deora, A. B., Kreitzer, G., Jacovina, A., and Hajjar, K. A. (2004) An annexin 2 phosphorylation switch mediates p11-dependent translocation of annexin 2 to the cell surface. *J. Biol. Chem*. **279**, 43411-43418

Dihazi, G. H. and Sinz, A. (2003) Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier Transform Ion-Cyclotron Resonance mass spectrometry. *Rapid Commun. Mass. Spectrom*. **17**, 2005-2014

Dole, M., Mack, L., and Hines, L. (1968) Molecular beams of macroions. *J. Chem. Phys*. **49**, 2240-2249

Dominguez, C., Boelens, R., and Bonvin, A. M. J. J. (2003) HADDOCK, A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc*. **125**, 1731 - 1737

Donato R. (2001) S100, A Multigenic Family of Calcium-Modulated Proteins of The EF-hand Type with intracellular and extracellular functional roles. *Int. J. Biochem. Cell Biol*. **33**, 637-668

Donato R. (1999) Functional roles of S100 proteins, calcium-binding proteins of the EF-hand type. *Biochim. Biophys. Acta* 1450, 191-231

Dreisewerd, K. (2003) The desorption process in MALDI. *Chem. Rev*. **103**, 395-426

Drum, C. L., Yan, S. Z., Bard, J., Shen, Y. Q., Lu, D., Soelaiman, S., Grabarek, Z., Bohm, A., and Tang, W. J. (2002) Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin. *Nature* **415**, 396-402

Eisenberg, D., Terwilliger, T. C., and Tsui, F. (1980) Structural studies of bee melittin. *Biophys. J*. **32**, 252-254

Engen, R. E. and Smith, D. L. (2001) Investigating protein structure and dynamics by hydrogen exchange MS. *Anal. Chem*. **73**, 256A-265A

Faure, A. V., Migne, C., Devilliers, G., and Ayala-Sanmartin J. (2002) Annexin 2 "secretion" accompanying exocytosis of chromaffin cells: possible mechanisms of annexin release. *Exp. Cell Res*. **276**, 79-89

Fenn, J. B., Mann, M., Meng C. K., Wong, S. F., and Whitehouse, C. M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64

Filipenko, N. R., MacLeod, T. J., Yoon, C.-S., and Waisman, D. M. (2004) Annexin A2 Is a Novel RNA-Binding Protein. *J. Biol. Chem*. **279**, 8723-8731

Fletcher, J. F. and Jiang, M. S. (1993) Possible mechanisms of action of cobra snake venom cardiotoxins and bee venom melittin. *Toxicon* **31**, 669-695

Gabel, F., Bicout, D., Lehnert, U., Tehei, M., Weik, M., and Zaccai, G. (2002) Protein dynamics studied by neutron scattering. *Quart. Rev. Biophys*. **35**, 327-367

Gamero-Castaño, M. and de la Mora, J. F. (2000) Mechanisms of electrospray ionization of singly and multiply charged salt cluster. *Analyt. Chim. Acta* **406**, 67-91

Gerke, V, and Weber, K. (1984) Identity of p36k phosphorylated upon rous sarcoma virus transformation with a protein purifies from brush borders, calcium-dependent binding to non-erythroid spectrin and F-Actin. *EMBO J*. **3**, 227-233

Gerke, V. and Moss, S. E. (2002) Annexins, from structure to function. *Physiol. Rev*. **82**, 331-371

Gerke V, Creutz CE, and Moss SE. (2005) Annexins, linking Ca^{2+} signalling to membrane dynamics. *Nat. Rev. Mol. Cell. Biol*. **6**, 449-461

Gest, J. E. and Salomon, Y. (1987) Inhibition by melittin and fluphenazine of melanotropin receptor function and adenylate cyclase in M2R melanoma cell membranes. *Endocrinology* **121**, 1766-1772

Glenney, J. R., and Track, B. F. (1985) Amino-terminal sequence of p36 and associated p10: identification of the site of tyrosine phosphorylation and homology with S-100. *Proc. Natl. Acad. Sci. USA* **82**, 7884-7888

Gray, J. J., Moughan, S. E., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A., and Baker, D. (2003) Protein-protein docking with simultaneous optimization of rigid body displacement and side chain conformations. *J. Mol. Biol.,* **331**, 281-299

Gross, J. H. Mass Spectrometry – A textbook. Springer Verlag Berlin Heidelberg, 2004

Hait, W. N., Cadman, E., Benz, C., Cole, J., Weiss, B. (1983) Inhibition of growth of L1210-leukemic cells by inhibitors of cyclic-nucleotide phosphodiesterase (PDE) and calmodulin (CAL), P. *Am. Assoc. Canc. Res*. **24**, 50

Hajjar, K. A. and Krishnan, S. (1999) Annexin II, a mediator of the plasmin/plasminogen activator system. *Trends Cardiovasc. Med*. **9**, 128-138

Hasan, A., Smith, D. L., and Smith, J.B. (2002) α-Crystallin regions affected by adenosine 5´-triphosphate identified by hydrogen-deuterium exchange *Biochemistry* **41**, 15876-15882

He, Z., Dunker, A.K., Wesson, C.R., and Trumble, W.R. (1993) Ca^{2+}-induced folding and aggregation of skeletal muscle sarcoplasmic reticulum calsequestrin. The involvement of trifluoperazine-binding site. *J. Biol. Chem*. **268**, 24635-24641

Heizmann CW, Fritz G, Schafer BW (1999) S100 proteins, structure, functions and pathology. *Biochim. Biophys. Acta*. **1450**, 191-231

Hering, J. A., Innocent, P. R., Haris, P.I. „Automatic amide I Frequency selection for rapid quantification of protein secondary structure from Fourier transform infrared spectra of proteins" Proteomics 2, **7,** 839-849 (2002)

Hermanson, G. T. „Bioconjugate Techniques" Academic Press, San Diego, CA, 1996

Horst, R., Wider, G., Fiaux, J., Bertelsen E. B., Horwich, A. L., and Wüthrich, K. (2006) Proton-proton Overhauser NMR spectroscopy with polypeptide chains in large structures. *Proc. Natl. Acad. Sci. USA* **103**, 15445-15550

Hurst, G. B., Lankford T. K., and Kennel, S. J. (2004) Mass spectrometric detection of affinity purified crosslinked peptides. *J. Am. Soc. Mass Spectrom*. 15, 832-839

Ihling, C., Schmidt, A., Kalkhof, S., Stingl, C., Mechtler, K., Haack, M., Schulz, D.M., Cooper, D.M.F., Beck-Sickinger, A.G., and Sinz, A. (2006)

Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A. (1992) Solution structure of a calmodulin-target peptide complex by multidimensional NMR, *Science* 256, 632-638

International Human Genome Sequencing Consortium „Initial sequencing and analysis of the human genome" *Nature* 409, 860-921 (2001)

Iribarne, J. V. and Thomson B. A. (1976) On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **64**, 2287-2294

Itakura, M. and Iio, T. (1992) Static and kinetic studies of calmodulin and melittin complex, *J. Biochem.* **112**, 183-191

James, P., Maeda, M., Fischer, R., Verma, A.K., Krebs, J., Penniston, J. T., and Carafoli, E. (1988) Identification and primary structure of a calmodulin binding domain of the Ca2+ pump of human erythrocytes, *J. Biol. Chem.* **263**, 2905-2910

Johnsson, N., Vandekerckhove, J., Van Damme, J., and Weber, K. (1986) Binding sites for Calcium, lipid and p11 on p36, the substrate of retroviral tyrosine-specific protein Kinases. *FEBS J.* **198**, 361-364

Johnsson, N., Mariott, G., and Weber, K. (1988) P36, the major cytoplasmic substrate of src tyrosine kinase, binds to its p11 regulatory subunit via a short amino-terminal amphiphatic helix. *EMBO J.*, **7**, 2435-2442

Johnsson, N., and Weber, K. (1990) Alkylation of cysteine 82 of P11 abolishes the complex formation with the tyrosine-protein Kinase Substrate P36 (Annexin 2, Calpactin I, Lipocortin 2). *J. Biol. Chem.* **265**, 14464-14468Jost M, Zeuschner D, Seemann J, Weber K, and Gerke V. (1997) Identification and Characterization of a Novel Type of Annexin-Membrane Interaction, Ca^{2+} is not required for the association of annexin with early endosomes. *J. Cell Sci.* **110**, 221-228

Kalkhof, S., Ihling, C., Mechtler, K., and Sinz, A. (2005) Chemical cross-linking and High-Performance Fourier Transform Ion Cyclotron mass spectrometry for protein interaction analysis: application to a calmodulin / target peptide complex. *Anal. Chem.* **77**, 495-503

Karas, M. and Hillenkamp, F. (1988) Laser desorption ionization of proteins with molecular masses exceeding 10000 Daltons *Anal. Chem.* **60,** 2299-2301

Karas, M. and Krüger, R. (2003) Ion formation in MALDI: the cluster ionization mechanism. *Chem. Rev.* **103**, 427-439

Kataoka, M., Head, J. F., Seaton, B. A., and Engelmann, D. M. (1989) Melittin binding causes a large calcium-dependent conformational change in calmodulin, *Proc. Natl. Acad. Sci. USA* **86**, 6944-6948

Kendrew, J. C. and Perutz, M. F. (1957) X-ray studies of compounds of biological interest. *Annu. Rev. Biochem.* **26**, 327-372

Knegtel, R. M. A., Boelens R., and Kaptein, R. (1994) Monte-Carlo docking of protein-DNA complexes, incorporation of DNA flexibility and experimental data. *Protein Eng.* **7**, 761-767

Knochenmuss, R. and Zenobi, R. (2003) MALDI ionization, the role of in-plume processes. *Chem. Rev.* **103**, 441-452

Kretsinger, R. H. and Nockolds, C. E. (1973) Carp muscle calcium-binding protein. II. Structure determination and general descriptin *J. Biol. Chem.* 248, 3313-3326.

Kretsinger, R. H. (1987) Cold Spring Harbor Symp. Quant. Biol. 52, 487-510

Kruppa G. H., Schoeniger J. S., and Young M. M.. A (2003) Top down approach to protein structural studies using chemical cross-linking and Fourier Transform mass spectrometry. *Rapid Commun. Mass Spectrom.* **17**, 155-162

Kube E, Becker T, Weber K, and Gerke V. (1992) Protein-protein interaction studied by site-directed mutagenesis. *J. Biol. Chem.* **267**, 14175-14182

Kundrotas, P. J. and Alexov E. (2007) PROTCOM, Searchable database of protein complexes enhanced with domain-domain structures. *Nucleic Acid Res.* **35**, D575-D579

Kurokawa, H., Osawa, M., Kurihara, H., Katayama, N., Tokumitsu, H., Swindells, M. B., Kainosho, M., and Ikura, M. (2001) Target-induced conformational adaptation of calmodulin revealed by the crystal structure of a complex with nematode Ca^{2+}/calmodulin-dependent kinase kinase peptide. *J. Mol. Biol.* **312**, 59-68

Lacroix, M., Rossi, V., Gaboriaud, C., Chevallier, S., Jaquinod, M., Thielens, N. M., Gagnon, J., and Arlaud, G. J. (1997) Structure and assembly of the catalytic region of human complement protease Clr, a three-dimensional model based on chemical cross-linking and homology modeling. *Biochemistry* **36**, 6270-6282

Lambert, O., Gerke, V., Bader, M.-F., Porte, F., and Brisson, A. (1997) Structural analysis of junctions formed between lipid membranes and several annexins by cryo-electron microscopy. *J. Mol. Biol.* **272**, 42-55

Lawrence, E. O. and Livingston, M. S. (1932) The production of high speed light ions without the use of high voltages. *Phys. Rev.* **40**, 19-35

Liemann S and Huber R. (1997) Three-dimensional structure of annexins. *Cell. Mol. Life Sci.* **53**, 516-521

Lee, D. B. N., Jamgotchian, N., Allen, S. G., Kan, F. W. K., and Hale, I. L. (2004) Annexin A2 heterotetramer, role in tight junction assembly. *Am. J. Physiol. Renal Physiol.* **287**, 481-491

Lee, G. L. and Hait, W. N. (1985) Inhibition of growth of C6 astrocytoma cells by inhibitors of calmodulin, *Life Sci.* **36**, 347-354.

Lehner, I., Niehof, M., and Borlak, J. (2003) An optimized method fort he isolation and identification of membrane proteins. *Electrophoresis,* **24**, 179-1808

Lei, H., Romeo, G., and Kazlauskas, A. (2004) Heat shock protein 90alpha-dependent translocation of annexin to the surface of endothelial cells modulates plasmin activity in the diabetic rat aorta. *Circ. Res.* **94**, 902-909

Lewit-Bentley A, Réty S, Sopkova-de Oliveira Santos J, and Gerke V. (2000) S-100 annexin complexes: some insights from structural studies. *Cell. Biol. Int.* **24**, 799-802

Loo, J. A. (1997) Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* **16**, 1-23

Lottspeich, F. and Zorbas (Editors), Bioanalytik, Spektrum Akademischer Verlag, Heidelberg, 1998

Malencik, D. A. and Anderson, S. R. (1988) Association of melittin with the isolated myosin light chains. *Biochemistry* **27**, 1941-1949

Mamyrin, B. A. (1994) Laser assisted reflectron time-of-flight mass spectrometry. *J. Mass Spectrom. Ion Proc.* **131**, 1-19

Maulet, Y. and Cox, J. A. (1983) Structural changes in melittin and calmodulin upon complex formation and their modulation by calcium. *Biochemistry* **22**, 5680-5686

Meador, W. E., Means, A. R., and Quiocho, F. A. (1992) Target enzyme recognition by calmodulin, 2.4 Å structure of a calmodulin-peptide complex. *Science* **257**, 1251-1255

Meador, W. E., Means, A. R., and Quiocho, F. A. (1993) Modulation of calmodulin plasticity in molecular recognition on the basis of X-ray structures. *Science* **262**, 1718-1721

Menke M, Ross M, Gerke V, and Steinem C. (2004) The molecular arrangement of membrane-bound annexin A2-S100A10 tetramer as revealed by scanning force microscopy. *Chem. Bio Chem.* **5**, 1003-1006

Menke, Manuela (2005) Visualisierung und Charakterisierung der molekularen Anordnung von Annexin A2 auf Lipidmembranen, http://www.opus-bayern.de/uni-regensburg/volltexte/2006/536/

Mickleburgh, I., Burtle, B., Hollas, H., Campbell, G., Chrzanowska-Lightowlers, Z., Vedeler, A., and Hesketh, J. (2005) Annexin A2 binds to the localization signal in the 3´ untranslated region of c-Myc mRNA. *FEBS. J.* **272**, 413-421

Mora, J. F. (2000) Electrospray ionization of large multiply charged species proceeds via Dole´s charged residue mechanism. *Analyt. Chim. Acta* **406**, 93-104

Morgan, R. O. and Fernandez, M. P. Distinct annexin subfamilies in plants and protists diverged prior to animal annexins and from a common ancestor. (1997) *J. Mol. Evol.* **44**, 178-188

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., and Olson, A. J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comp. Chem.* **19**, 1639-1662

Moss SE and Morgan RO. (2004) The Annexins. *Genome Biol.* **5**, 219

Müller, D. R., Schindler, P., Towbin, H., Wirth, U., Voshol, H., Hoving, S., and Steinmetz, M. O. (2001) A new tool in mass spectrometric protein interaction analysis. *Anal. Chem.* **73**, 1927-1934

Nakata, T., Sobue, K., and Hirokawa, N. (1990) Conformational change and localization of calpactin I complex involved in exocytosis as revealed by quick-freeze deep-etch electron microscopy. *J. Cell Biol.* **110**, 13-25

Navaza, J. (1994) AMoRe: an Automated package for Molecular Replacement. *Acta Crystallogr. A* **50**, 157-163

Nguyen, H. T. *et al.* (2006) Proteomic characterization of lipid raft markers from the intestinal brush border. *Biochem. Biophys. Res. Commun.* **342**, 236-244

Nilges, M., Gronenborn, A. M., Brünger, A. T. and Clore, G. M. (1988) Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2, *Prot. Eng.* **2**, 27-38

Novak, P., Young M. M., Schoeniger J. S., and Kruppa G. H. (2003) A Top-Down Approach to Protein Structure Studies Using Chemical Cross-Linking and Fourier Transform Mass Spectrometry. *Eur. J. Mass Spectrom.* **9**, 623-631

Novak, P., Haskins, W. E., Ayson, M. J., Jacobsen, R. B., Schoeniger, J. S., Leavell, M. D., Young, M. M., and Kruppa, G. H. (2005) Unambiguous assignment of intramolecular chemical cross-links in modified mammalian membrane proteins by Fourier Transform-Tandem mass spectrometry. *Anal. Chem.* **77**, 5101-5106

Olmea, O., Straus, C. E., and Ortiz, A. R. (2002) Mammoth (Matching Molecular Models Obtained from Theory): an automated method for model comparison. *Protein Sci.* **11**: 2606-2621 O´Neil, K. T. and DeGrado, W. F. (1990) How calmodulin binds its targets, sequence independent recognition of amphiphilic α-helices, *Trends Biochem. Sci.* **15**, 59-64

Orsolic, N., Sver, L., Verstovsek, S., Terzic, S., and Basic, I. (2003) Inhibition of mammary carcinoma cell proliferation in vitro and tumor growth in vivo by bee venom. *Toxicon* **41**, 861-870

Osawa, M., Tokumitsu, H., Swindells, M. B., Kurihara, H., Orita, M., Shibanuma, T., Furuya, T., and Ikura, M. (1999) A novel target recognition revealed by calmodulin in complex with Ca^{2+}-calmodulin-dependent kinase kinase, *Nat. Struct. Biol.* **6**, 819-824

Persechini, A. and Kretsinger, R. H. (1988) The central helix of calmodulin functions as a flexible tether, *J. Biol. Chem.* **263**, 12175-12178

Rand, J. H. (2000) The Annexinopathies: a new category of diseases. *Biochim. Biophys. Acta* **1498**, 169-173

Raynal, P. and Pollard, H. B. (1994) Annexins: the problem of assessing the biological role for a gene family of multifunctional calcium- and phospholipid-binding proteins. *Biochim. Biophys. Acta* **1197**, 63-93

Raynor, R. J., Zheng, B., and Kuo, J. F. (1991) Membrane interactions of amphiphilic polypeptides mastoparan, melittin, polymyxin B, and cardiotoxin. Differential inhibition of protein kinase C, Ca^{2+}/calmodulin-dependent protein kinase II and synaptosomal membrane Na,K-ATPase, and Na^+ pump and differentiation of HL60 cells. *J. Biol. Chem.* **266**, 753-2758

Rescher, U. and Gerke, V. (2004) Annexins: unique membrane-binding proteins with diverse functions. *J. Cell Sci.* **117**, 2631-2639

Réty, S., Sopkova, J., Renouard, M., Tabares, S., Osterloh, D., Gerke, V., Russo-Marie, F., and Lewit-Bentley, A. (1999) The crystal structure of a complex of p11 with the annexin II N-terminal peptide. *Nat. Struct. Biol.* **6**, 89-95

Rosengarth, A. and Luecke, H. (2004) Annexin A2 - Does it induce membrane aggregation by a new multimeric state of the protein? *Annexins* **1**, 129-136

Scaloni, A., Miraglia, N., Orrù, S., Amodeo, P., Motta, A., Marino, G., and Pucci, P. (1998) Topology of the calmodulin-melittin complex, *J. Mol. Biol.* **277**, 945-958

Schmidt, A., Kalkhof, S., Ihling, C., Cooper, D. M. F., and Sinz A. (2005) Mapping protein interfaces by chemical cross-linking and FTICR mass spectrometry: application to a calmodulin/adenylyl cyclase 8 peptide complex. *Eur. J. Mass Spectrom.* **11**, 525-534

Schulz DM, Ihling C, Clore GM, and Sinz A. (2004) Mapping the topology and determination of a low-resolution three-dimensional structure of the calmodulin-melittin complex by chemical cross-linking and high-resolution FTICRMS: direct demonstration of multiple binding modes. *Biochemistry* **43**, 4703-4715

Schulz DM, Kalkhof, S., Schmidt, A., Ihling, C., Stingl, C., Mechtler, K., Zschörnig, O., and Sinz, A. (2007) Annexin A2 / p11 Interaction: New Insights Into Annexin A Tetramer Structure by Chemical Cross-Linking, High-Resolution Mass Spectrometry, and Computational Modeling, *PROTEINS: Structure, Function, and Bioinformatics* **69**, 254-269

Schumacher, M. A., Rivard, A. F., Bächinger, H. P., and Adelman, J. P. (2001) Structure of the gating domain of a Ca2+-activated K^+ channel complexed with calmodulin, *Nature* **410**, 1120-1124

Schwieters, C. D. and Clore, G. M. (2001) Internal coordinates for molecular dynamics and minimization in structure determination and refinement, *J. Magn. Reson.* **152**, 288-302

Sinz, A. (2003) Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *J. Mass Spectrom.* **38**, 1225-1237

Sinz, A., Kalkhof, S., and Ihling, C. (2005) Mapping protein interfaces by a trifunctional cross-linker combined with MALDI-TOF and ESI-FTICR mass spectrometry. *J. Am. Soc. Mass Spectrom.* **16**, 1921-1931

Sinz, A. (2006b) Isotope-labeled cross-linkers and Fourier Transform Ion Cyclotron Resonance mass spectrometry for structural analysis of a protein / peptide complex, *J. Am. Soc. Mass Spectrom.* **17**, 1100-1113

Sinz, A. (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev*, **25**, 663-682

Sopkova-de Oliveira Santos J, Oling FK, Réty S, Brisson A, Smith JC, and Lewit-Bentley A. (2000) S100 protein-annexin interactions: a model of the (Anx2-p11)$_2$ heterotetramer complex. *Biochim. Biophys. Acta* **1498**, 181-191

Stephens, W. (1946) Phys. Rev. 69, 691

Stryer, L. Biochemistry W. H. Freeman and Company, New York, 1995

Suckau, D., Mak, M., and Przybylski, M. (1992) Protein surface topology-probing by selective chemical modification and mass-spectrometric peptide mapping. *Proc. Natl. Acad. Sci. USA* **89**, 5630-5634

Swaim, C. L., Smith, J. B., and Smith, D. L. (2004) Unexpected products from the reaction of the synthetic cross-linker 3,3´-dithiobis(sulfosuccinimidyl propionate), DTSSP with peptides. *J. Am. Soc. Mass Spectrom.* **15**, 736-749Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida , Y., Yoshida, T., and Matsuo, T. (1988) Protein and polymer analyses up to *m/z* 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **2**, 151-153

Taverner, T., Hall, N. E., O`Hair, R. A. J., and Simpson, R. J. (2002) Characterization of an antagonist interleukin-6 dimer by stable isotope labeling, cross-linking, and mass spectrometry. *J. Biol. Chem.* **277**, 46487-46492

Terwilliger, T. C. and Eisenberg, D. (1982) The structure of melittin. II. Interpretation of the structure. *J. Biol. Chem.* **257**, 6016-6022

Thomson, J. J. (1906) Carriers of negative electricity (Nobel lecture), www.nobel.se/physics/laureates/1906/thomson-lecture.html

Tomonaga, T., Matsushita, K., Yamaguchi, S., Oh-Ishi, M., Kodera, Y., Maeda, T., Shimada, H., Ochiai, T., and Nomura, F. (2004) Identification of altered protein expression and post-translational modifications in primary colorectal cancer by using agarose two-dimensional gel electrophoresis. *Clin. Cancer Res.* **10**, 2007-2014

Trester-Zedlitz M, Kamada K, Burley SK, Fenyö D, Chait BT, and Muir TW. (2003) A modular cross-linking approach for exploring protein interactions. *J. Am. Chem. Soc.* **125**, 2416-2425

Tung, C. S., Walsh, D. A., and Trewhella J. (2002) A structural model of the catalytic subunit-regulatory subunit dimeric complex of the cAMP-dependent protein kinase. *J. Biol. Chem.* **277**, 12423-12431

van Dijk, A. D. J., Boelens, R., and Bonvin, A. M. J. J. (2005) Data-driven docking for the study of biomolecular complexes. *FEBS J.* **272**, 293-312

Venter, J.C. *et al.* (2001) The sequence of the human genome, *Science* **291**, 1304-1351

Vetter, S. W. and Leclerc, E. (2003) Novel aspects of calmodulin target recognition and activation, *Eur. J. Biochem.* **270**, 404-414

von Heijne, G. (2006) Membrane-protein topology. *Nature Rev. Mol. Cell Biol.* **7**, 909-918

Waisman DM. (1995) Annexin II tetramer: structure and function. *Mol. Cell Biochem.* **149/150**, 301-322

Wang, G., Louis, J. M., Sondej, M., Seok, Y. J., Peterkofsky, A. and Clore, G. M. (2000) Solution structure of the phosphoryl transfer complex between the signal transducing proteins HPr and IIAGlucose and HPr of the Escherichia coli phosphoenolpyruvate,sugar phosphotransferase system, *EMBO J.* **19**, 5635-5649

Watterson, D. M., Sharief, F., and Vanaman, T. C. (1980) The complete amino acid sequence of the Ca^{2+}-dependent modulator protein (calmodulin) of bovine brain. *J. Biol. Chem.* **255**, 962-975

Weber, P. J. and Beck-Sickinger, A. G. (1997) Comparison of the photochemical behavior of four different photoactivatable probes. *J. Pept. Res*. **49**, 375-383

Wiley, W, C, and McLaren, I. H. (1955) Time-of-flight mass spectrometer with improved resolution. *Rev. Sci. Instrum.* **26**, 1150-1157

Wu T, Angus W, Yao X-L, Logun C, and Shelhamer JH. (1997) P11, a unique member of the S100 family of calcium-binding proteins interacts with and inhibits the activity of the 85-kDa cytosolic phospholipase A2. *J. Biol. Chem*., **272**, 17145-17153

Wüthrich, K. (2003) NMR-Untersuchungen von Struktur und Funktion biologischer Makromoleküle (Nobel lecture), *Angew. Chemie* **115**, 3462-3486

Yuan, T., Weljie, A. M., and Vogel, H. J. (1998) Tryptophan fluorescence quenching by methionine and selenomethionine residues of calmodulin, Orientation of peptide and protein binding. *Biochemistry* **37**, 3187-3195

Zhang, Z. and Smith, D. L. (1993) Determination of amide hydrogen exchange by mass spectrometry, a new tool for protein structure elucidation. *Protein Sci.* **2**, 522-531

Publications and Presentations

Research Articles

Schulz, D. M., Kalkhof, S., Schmidt, A., Ihling, C., Stingl, C., Mechtler, K., Zschörnig, O., and Sinz, A., Annexin A2 / p11 Interaction: New Insights Into Annexin A Tetramer Structure by Chemical Cross-Linking, High-Resolution Mass Spectrometry, and Computational Modeling, *PROTEINS: Structure, Function, and Bioinformatics* 69, 254-269 (2007)

Watanabe, T., **Schulz, D.**, Morisseau, C., and Hammock, B.D., High-Throughput Pharmacokinetic Method: Cassette Dosing in Mice Associated with Minuscule Serial Bleedings and LC-MS/MS Analysis, *Analytica Chimica Acta* 559, 37-44 (2006)

Ihling, C., Schmidt, A., Kalkhof, S., Stingl, C., Mechtler, K., Haack, M., **Schulz, D.M**., Cooper, D.M.F., Beck-Sickinger, A.G., and Sinz, A., Isotope-Labeled Cross-Linkers and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for Structural Analysis of a Protein / Peptide Complex, *Journal of the American Society for Mass Spectrometry* 17, 1100-1113 (2006)

Schulz, D.M., Ihling, C., Clore, G.M., and Sinz, A., Mapping the Topology and Determination of a Low-Resolution Three-Dimensional Structure of the Calmodulin-Melittin Complex by Chemical Cross-Linking and High-Resolution FTICR Mass Spectrometry: Direct Demonstration of Multiple Binding Modes, *Biochemistry* 43, 4703-4715 (2004)

Reviews

Schulz, D.M. und Sinz, A., Chemisches Cross-Linking und Massenspektrometrie zur strukturellen und funktionellen Charakterisierung von Proteinen, *BIOSpektrum* 10, 45-48 (2004)

Selected Poster Presentations

Schulz, D.M., Schmidt, A., Zschörnig, O., Kalkhof, S., Ihling, C., and Sinz, A., Chemical Cross-Linking, Mass Spectrometry, and Computational Methods for the Determination of a Low-Resolution Structure of the Annexin A2 / p11 Heterotetramer, *Deutsche Gesellschaft für Massenspektrometrie, 39. Diskussionstagung*, 05.-08.03.2006, Mainz

Schulz, D.M., Schmidt, A., Zschörnig, O., Kalkhof, S., Ihling, C., and Sinz, A., Purification and Characterization of the Tetrameric Complex Between Annexin II and p11 from Porcine Small Intestine, *Deutsche Gesellschaft für Massenspektrometrie, 38. Diskussionstagung*, 06.-09.03.2005, Rostock

Schulz, D.M., Ihling, C., Clore, G.M., and Sinz, A., Elucidation of Low-Resolution Three-Dimensional Structures of the Calmodulin / Melittin Complex by Chemical Cross-Linking and FTICR Mass Spectrometry, *Proceedings of the 52nd ASMS Conference on Mass Spectrometry and Allied Topics* (2004), Nashville, TN, USA

Schulz, D.M, Ihling, C., Clore, G.M., and Sinz, A., Mapping Protein Interfaces by Chemical Cross-Linking and FTICR Mass Spectrometry: Application to the Calmodulin / Melittin Complex, *Deutsche Gesellschaft für Massenspektrometrie, 37. Diskussionstagung*, 07.-10.03.2004, Leipzig

Schulz, D.M, Ihling, C., Clore, G.M., and Sinz, A., Mapping Protein Interfaces by Chemical Cross-Linking and FTICR Mass Spectrometry: Application to the Calmodulin / Melittin Complex, *2nd Leipzig Research Festival for Life Sciences*, 24.10.2003, Leipzig (**Posterpreis**)

Schulz, D.M, Ihling, C., Clore, G.M., and Sinz, A., Mapping Protein Interfaces by Chemical Cross-Linking and FTICR Mass Spectrometry: Application to the Calmodulin / Melittin Complex, *Proteomic Forum,* 14.-17.09.2003, München

Schulz, D.M. and Sinz, A., Investigation of Protein Interfaces by Chemical Cross-Linking and Mass Spectrometry, *Deutsche Gesellschaft für Massenspektrometrie, 36. Diskussionstagung*, 09.-12.03.2003, Münster

Curriculum Vitae

Name	Daniela Michaela Schulz
Geburtsdatum	28. Juli 1978
Geburtsort	Nürnberg
Familienstand	ledig

Schulbildung

1984-1988	Adalbert-Stifter-Grundschule in Nürnberg
1988-1991	Gymnasium der Maria-Ward-Schule in Nürnberg
1991-1997	Pirckheimer-Gymnasium in Nürnberg
1997	Abitur

Studium

1997 -2002	Diplomstudium Agrarwissenschaften (Schwerpunkt Agrarbiologie) an der Technischen Universität München – Weihenstephan
2002	Aufenthalt an der University of California, Davis
Oktober 2002	Diplomarbeit zum Thema „Metabolism of Soluble Epoxide Hydrolase Inhibitors"

Promotion

Oktober 2002	Beginn der Promotion an der Universität Leipzig – Biotechnologisch-Biomedizinisches Zentrum, Institut für Analytische Chemie

Halle (Saale), 14. Februar 2007

Danksagung

Selbständigkeitserklärung

Ich versichere hiermit, die vorliegende Dissertation selbstständig und ohne unerlaubte fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die im Literaturverzeichnis angeführten Quellen genutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen als solche gekennzeichnet.

Bei der Auswahl und Auswertung des Materials, bei der Herstellung des Manuskripts, sowie bei der geistigen Herstellung der vorgelegten Arbeit waren keine anderen Personen beteiligt. Insbesondere wurde weder die Hilfe eines Promotionsberaters in Anspruch genommen, noch haben Dritte von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die vorgelegte Arbeit ist weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt worden.

Ich habe keine früheren erfolglosen Promotionsversuche unternommen.

---   Halle (Saale), 14. Februar 2007

Daniela Schulz